REMARKS

Claims 1, 17, 19, 20, 35, 42, 47, 78, 93 and 106-126 are pending in the subject application. Claims 2, 3, 6, 11-15, and 21 have been cancelled without prejudice in this amendment. Claims 17, 42, 47, 78, 93, 106-110, 112, and 115 have been amended in this amendment. Applicant reserves the right to pursue the previously claimed subject matter in one or more continuing applications. Claims 118-126 have been added. No new matter has been added. Amendments to claims 17, 42, 47, and 78 correct minor grammatical or typographical errors. Support for the amendments to claims 93, 106-108, 112, and 115 and for claims 118-126 is found throughout the specification, including original and previously presented claims. For example, support for the amendment to claims 93, 106, and 107 is found at least in previous claim 108 and on p. 43, lines 6-12; support for amendment to claim 108 is found at least on p. 48 (first full paragraph); claims 109-110 were amended to depend on claim 93 rather than claim 108, consistent with the amendment of claim 93 to recite "increase" rather than "change"; support for amendment to claim 112 is found at least on p. 46 (first full paragraph); support for amendment to claim 115 is found at least on p. 47, lines 19-21. Support for claims 118-120 is found at least on p. 21 (first paragraph), p. 46 (first full paragraph), and p. 50 (first full paragraph). Support for claims 121-123 is found at least in the paragraph bridging pp. 45-46, paragraph bridging pp. 48-49, p. 48 (first full paragraph), and p. 49 (first full paragraph). Support for claims 124 and 125 is found at least on p. 11, lines 12-13. Support for claim 126 is found at least in original claim 94 and paragraph bridging pp. 46-47.

Elections/Restrictions

In the Office Action claims 113-117 were withdrawn as allegedly being drawn to methods distinct from that of the elected invention. For each of the following reasons, Applicant respectfully submits that the Office has not met the burden of establishing that restriction between claim 93 and claims 113-117 is proper and respectfully requests that claims 113-117 be rejoined and examined.

First, as set forth in MPEP 806.05(j), restriction between related process inventions is only proper if the inventions as claimed do not overlap in scope. Claims 113-117 depend on claim 93 and therefore include each step recited in claim 93. Applicant respectfully submits that since claims 113-114 require each step of claim 93, claims 113-114 must fall within the scope of claim 93. Applicant therefore respectfully disagrees with the statement that claims 113-114 are "not

commensurate in scope with claim 93". Similarly, claims 115-116 require each step of claim 93 and thus also fall within the scope of claim 93. See MPEP 806.05(j), which indicates that restriction between related process inventions is only proper if the inventions are mutually exclusive, such that there is no process that would infringe both inventions. Applicant respectfully submits that the process of any of claims 113-117 would infringe claim 93. Therefore, restriction is not proper.

Second, Applicant respectfully notes that in accordance with MPEP 803, "If the search and examination of all the claims in an application can be made without serious burden, the examiner must examine them on the merits, even though they include claims to independent or distinct inventions." Applicant respectfully submits that the Office has not established that examining claims 113-117 together with claims 93 and 106-112 would result in a serious burden on the Examiner. Since each of claims 113-117 recites every step of claim 93, and an additional step (i.e., a step of "assessing" or "administering"), Applicant respectfully submits that the additional burden, if any, imposed by examining claims 113-117 once having examined claim 93 is slight. Applicant further respectfully submits that the Office Action did not set forth an appropriate explanation of separate classification, separate status in the art, or different field of search as defined in, and required by, MPEP 808.02, for imposing a restriction between claims 93 and 113-114 or between claim 93 and claims 115-117.

Third, in response to the Office's contention that claims 115-117 are drawn to a method of treating a mammal with a disease (Office Action, p. 2, last paragraph), Applicant respectfully submits that claims 115-117 recite a method that includes steps (a) and (b) of claim 93 and further comprises a step of administering an agent to a mammalian organism. There is no requirement in these claims of treating a disease. For example, after identifying an agent according to the method of claim 93, the step of administering set forth in claim 115 might be performed as a secondary screen, e.g., to assess whether the agent has a particular biological activity in vivo. For example, as described in the specification, "In other embodiments, a secondary screening step is performed on the agent. In a specific embodiment, the agent is tested in additional assays for its effects on mitochondrial cell number or a mitochondrial function, such as coupled oxygen consumption.... Agents identified using the methods of the present invention may also be tested in model systems for their efficacy in inducing the desired biological response or in treating disorders." (p. 47, lines 6-21). The specification further states, "Toxicity and therapeutic efficacy of the agents and

compositions of the present invention can be determined by standard pharmaceutical procedures in cell cultures or experimental animals." (p. 67, lines 23-28). Applicant respectfully submits that the agent recited in claims 115-17 might be administered to assess the effect(s) of the agent or to assess whether the agent increases a biological response, as recited in claims 113-114. Applicant therefore submits that it would be appropriate to examine claims 115-117 together with claims 113-114.

In summary, for each of the above reasons Applicant respectfully submits that claims 113-117 were improperly withdrawn and respectfully requests that they be rejoined and examined.

Specification

The disclosure was objected to because the legends of Figures 3, 4, and 5 allegedly do not match the figures. Replacement Figures 3, 4, and 5, in which the panels are labeled in accordance with the figure legends, are submitted herewith.

Rejections under 35 U.S.C. 112, first paragraph

Claims 93 and 106-112 are rejected as allegedly lacking enablement. The Office maintains that the specification and art provide no guidance that changing the expression pattern of any of the OXPHOS-CR genes can be used to treat diabetes or a mitochondrial disease or disorder. The Office concludes that decreased OXPHOS-CR gene expression appears "correlative rather than causative of any disease or disorder" and that it would be undue experimentation for an artisan to use agents identified in the claimed screen for treating diabetes or a mitochondrial disease or disorder (Office Action, paragraph bridging pp. 5-6). The Office also maintains that it would be undue experimentation to practice the claimed invention without knowing which combination of OXPHOS-CR genes to use (Office Action, p. 6, first full paragraph). The Office also comments that the specification does not teach that OXPHOS-CR genes are downregulated in animals other than diabetic humans such that upregulation of OXPHOS-CR genes in diseased animals will result in treatment of these animals (Office Action, paragraph bridging pp. 6-7). Applicants respectfully disagree and request reconsideration for at least the following reasons.

First, with respect to the Examiner's comment that the artisan would not know which OXPHOS-CR genes to use, Applicant respectfully submits that, at least for the reason that expression of OXPHOS-CR genes is tightly co-regulated across multiple tissues (specification, p. 81, line 7), the choice of a particular pair of OXPHOS-CR genes is not critical to the claimed

method. One of skill in the art could select any of a large number of different pairs of OXPHOS-CR genes whose expression change in response to an agent is representative of the OXPHOS-CR genes set as a whole.

Second, Applicant respectfully submits that decreased OXPHOS-CR gene expression does not need to be "causative" of diabetes or a mitochondrial disease or disorder in order for compounds identified in the claimed screen to be effective for treating such diseases. Such a relationship is not required by the claims, and Applicant respectfully submits that there are many examples of effective drugs that act on a gene or gene product that has not been established as "causative" of the disease for which the drug is administered. The Office acknowledges that the specification teaches that OXPHOS-CR genes are downregulated in DM2 and IGT patients and that decreased OXPHOS-CR gene expression correlates with these conditions (Office Action, paragraph bridging pp. 5-6). Applicant notes that the specification also discloses additional discoveries regarding OXPHOS-CR genes and their physiological significance and regulation that also support the enablement of the claimed screening method (see, e.g., Examples 3-6). Applicant submits that in accordance with the instant invention, the correlative relationship provides sufficient basis for a screen to identify effective agents, and one of skill in the art reading the present specification would recognize that an increase in OXPHOS-CR gene expression would reflect an alteration in the physiological state of a subject suffering from such disease towards a more normal state. A compound capable of causing such alteration would therefore be expected to be of therapeutic use. Expressed another way, Applicant submits that increased expression of OXPHOS-CR genes in cells contacted with a compound serves as a "marker" indicating that the compound shifts the physiological state of the cell away from the state found in disorders characterized by glucose intolerance, insulin resistance or reduced mitochondrial function and towards a more normal state.

Third, Applicant respectfully submits that there is substantial evidence supporting the contention that agents that increase OXPHOS-CR gene expression (as would be identified using the claimed methods) will be effective in treating diabetes and/or a mitochondrial disease or disorder. Applicant and his collaborators screened a small molecule library and successfully identified at least two compounds that increase OXPHOS-CR expression and that had been previously reported to have anti-diabetic effects. Specifically, gene expression–based high-throughput screening was performed to profile transcripts in differentiated C2C12 myotubes following 48-h treatment with each of 2,490 compounds (see Wagner BK, Kitami T, Gilbert TJ, Peck D, Ramanathan A, Schreiber

SL, Golub TR, Mootha VK, "Large-scale chemical dissection of mitochondrial function", Nat Biotechnol., 26(3):343-51, 2008, provided herein as Exhibit A, hereinafter "Wagner et al."). The research sought to identify compounds that promote OXPHOS gene expression while reducing ROS levels (see p. 344, the first paragraph on p. 346, and the sentence bridging pp. 346-347 of Exhibit A). As described in Wagner et al., "a list of highly co-regulated OXPHOS genes that are coordinately expressed across tissues and are downstream of the PGC-1 transcriptional coactivator was used" in their gene expression profiling (third full paragraph, right column, p. 349). These genes are the OXPHOS-CR genes of the present invention. [For evidence indicating that the highly co-regulated OXPHOS genes profiled in Wagner et al. are the OXPHOS-CR genes of the instant specification, please see Wagner et al., third full paragraph, right column, p. 349, citing reference 4 on which Applicant is the first author (Mootha VK, et al., PGC-1-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes, Nature Genetics, 34, 267 – 273, 2003, enclosed herein as Exhibit B). Reference 4 discloses the identification of OXPHOS-CR genes on p. 269 and lists them in the legend to Figure 3. Note that the list of genes in Figure 3 is identical to the list of OXPHOS-CR genes disclosed in the instant application on p. 46, second full paragraph.]

The screen identified microtubule inhibitors, such as mebendazole and deoxysappanone B, as compounds that increase OXPHOS-CR gene expression (Wagner et al., p. 347, first full paragraph). Notably, both of these compounds had been previously reported to have anti-diabetic effects. As disclosed on p. 349, "Type 1 and type 2 diabetics treated for parasitic worm infections with mebendazole unexpectedly showed improvement both in fasting blood glucose levels and nonesterified fatty acid metabolism. Blood glucose levels improved without weight loss, suggesting that the improvement in diabetes was not due to a toxic side effect of mebendazole. Moreover, deoxysappanone B, a natural product found in sappan wood, has been reported to be an antidiabetic component of Chinese herbal medicine... (Wagner et al, p. 349, right column, last full paragraph)." Applicant respectfully submits that the results reported in Wagner et al clearly establish that the claimed screening method is enabled for identifying compounds useful to treat diabetes or a mitochondrial disease or disorder. Indeed, the fact that at least two compounds with previously documented anti-diabetic effects were identified in screening a library of less than 2,500 compounds demonstrates the power of the instantly claimed methods.

Furthermore, the art provides examples of compounds that are in clinical use for treating various disorders characterized by glucose intolerance, insulin resistance or reduced mitochondrial function and that cause upregulation of OXPHOS-CR genes and would therefore be identified using the claimed methods. For example, pioglitazone is a compound that lowers insulin resistance and was approved by the US Food & Drug Administration in 1999 for treatment of DM2. Pioglitazone also finds use in treatment of other disorders characterized by insulin resistance such as polycystic ovary syndrome (PCOS). See, e.g., the first two paragraphs of the introduction of Skov, V., et al., "Pioglitazone enhances mitochondrial biogenesis and ribosomal protein biosynthesis in skeletal muscle in polycystic ovary syndrome", PLoS ONE. Jun 18;3(6):e2466, 2008 (hereinafter Skov et al., provided herein as Exhibit C). Skov et al. studied the effect of pioglitazone on gene expression in skeletal muscle of 10 obese women with polycystic ovary syndrome (PCOS) using global transcriptional profiling and then used gene set enrichment analysis (GSEA; a methodology disclosed in the instant specification, see Example 2, pp. 78-79) to identify the 15 most upregulated gene sets analyzed. Skov et al. showed that treatment with pioglitazone results in increased OXPHOS gene expression and report that their "major finding is that the expression of genes representing OXPHOS pathways is upregulated in skeletal muscle of PCOS patients together with an increase in insulin sensitivity in response to 16 weeks treatment with pioglitazone." (see Skov et al., Discussion). As disclosed in the instant specification, the bulk of the statistical signal indicative of downregulation of OXPHOS genes observed in DM2 is accounted for by OXPHOS-CR genes (Example 4, paragraph bridging pages 80-81). Applicant submits that, absent evidence to the contrary, one of skill in the art would conclude that the upregulation of OXPHOS genes observed by Skov et al. following pioglitazone treatment reflects upregulated expression of at least some of the OXPHOS-CR subset. Thus it is clear that the claimed methods would identify pioglitazone, a drug in clinical use to treat disorders characterized by insulin resistance.

With respect to the portion of the rejection alleging that the claims lack enablement because the specification allegedly "does not provide guidance that other animals can be treated with agents obtained from the claimed screen using human OXPHOS-CR genes", Applicant respectfully disagrees for at least the following reasons.

First, Applicant respectfully submits that among claims 93 and 106-112, only claims 109 and 111 recite that the agent is a potential agent for the treatment of a disorder that is characterized

by glucose intolerance, insulin resistance or reduced mitochondrial function. Therefore, the question of whether the other animals can be treated with agents identified from the claimed screen is not relevant to enablement of claims 93, 106-108, 110, and 112.

Second, Applicant respectfully submits that, "In order to make a rejection, the examiner has the initial burden to establish a reasonable basis to question the enablement provided for the claimed invention. *In re Wright*, 999 F.2d 1557, 1562, 27 USPQ2d 1510, 1513 (Fed. Cir. 1993). Applicant respectfully submits that it is well accepted that mammals such as mice and rats share many aspects of basic physiology with humans and serve as useful models human disease. Indeed the premise that such mammals share similarities with humans provides much of the motivation for conducting research on these animals in the first place. The specification teaches that "agents may also be tested for their efficacy in treating diabetes by using a non-obese diabetic (NOD) mouse. The successful use of this animal model in diabetic drug discovery is reported in the literature..." and cites numerous references describing such use (p. 47, lines 19-29). Furthermore, Applicant submits that there are rat models of diabetes used in the art, and it is well known that cats spontaneously develop diabetes resembling the human disease. The specification also teaches that "A 'patient' or 'subject' to be treated by the method of the invention can mean either a human *or non-human animal*, preferably a mammal." (p. 11, lines 12-13, emphasis added).

Furthermore, Applicant demonstrated important similarities between expression pattern and regulation of OXPHOS-CR genes in mice and humans that support the contention that agents that modulate expression of these genes would have similar useful effects in different species. As described on p. 71, lines 11-17 and in Example 4, analysis of the expression of these OXPHOS genes in a public atlas of mouse gene expression, showed that 2/3 of all OXPHOS genes (i.e., the OXPHOS-CR genes) are tightly co-regulated across all tissues examined, and that they are highly expressed at the major sites of insulin mediated glucose uptake in mice (brown fat, heart, and skeletal muscle. The strong correlation in expression of the OXPHOS-CR genes and their coordinated downregulation in human diabetic muscle led Applicant to hypothesize that the transcriptional co-activator PPARGC1 (PGC-1α) was responsible for this transcriptional co-regulation. Applicant observed that mean levels of PGC-1 α transcript were similarly decreased (-20%) in the diabetic muscle and noted that the promoters of several of the OXPHOS-CR genes have been reported to contain binding sites for nuclear respiratory factor 1, a transcription factor co-

activated by PGC-1α. Applicant infected mouse muscle cell lines with PPARGC1 and demonstrated that the OXPHOS-CR genes are specifically induced in a time-dependent manner over a three day period (p. 71, lines 11-20). Furthermore, as described in Example 16, Applicant showed that Errα and Gaba motifs are evolutionarily conserved between mice and humans and are enriched upstream of OXPHOS genes (pp. 94-95). Applicant thus concluded that PGC-1α, Errα and Gabp regulate the expression of level of OXPHOS genes in both mouse and humans. Applicant respectfully submits that the similar expression pattern of OXPHOS-CR genes, with high expression at major sites of insulin mediated glucose uptake in mice and humans, as well as the evidence for similar regulation via PGC-1α, Errα and Gaba, would lead one of skill in the art to conclude that the OXPHOS-CR genes play similar roles in humans and other mammals and that agents that increase OXPHOS-CR gene expression will likewise be of use to treat both humans and other mammals suffering from diabetes or mitochondrial disorders.

As stated by the court in *In re Marzocchi*, "it is incumbent upon the Patent Office, whenever a rejection on this basis is made, to explain why it doubts the truth or accuracy of any statement in a supporting disclosure and to back up assertions of its own with acceptable evidence or reasoning which is inconsistent with the contested statement. Otherwise, there would be no need for the applicant to go to the trouble and expense of supporting his presumptively accurate disclosure." *In re Marzocchi*, 439 F.2d 220, 224, 169 USPQ 367, 370 (CCPA 1971). Hoshikawa et al. relates solely to changes in gene expression in rats and mice exposed to hypoxia and does not teach or suggest that OXPHOS-CR genes would not be downregulated in animals other than humans. The single example of alleged differences between gene expression in rats and mice exposed to hypoxia reported by Hoshikawa et al. does not provide a reasonable basis to overcome the presumption that the claims are enabled, particularly in view of the evidence described above.

With respect to the fact that original claim 93 encompassed identifying agents that decrease OXPHOS-CR gene expression (Office Action, p. 7), while Applicant respectfully disagrees that this is a proper basis for rejecting the claims, solely in order to advance prosecution claim 93 has been amended to recite, "determining whether the expression of at least two OXPHOS-CR gene products show a coordinate increase in the test cell compared to an appropriate control, wherein a coordinate increase in the expression of the OXPHOS-CR gene products indicates that the agent regulates the expression of OXPHOS-CR genes", thereby obviating this ground for rejection.

In summary, Applicant respectfully submits that the claims are enabled and respectfully requests reconsideration and withdrawal of the rejection.

Rejection under 35 U.S.C. 102

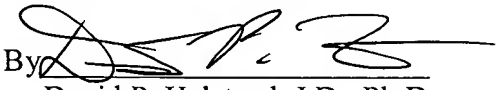
Claims 93, 106, and 107 are rejected under 35 U.S.C. 102(b) as allegedly being anticipated by Burke et al., 1997, The Journal of Biological Chemistry, 272: 14705-14712. As noted above, claim 93 has been amended to recite that the expression of at least two OXPHOS-CR gene products show a coordinate increase in the test cell as compared to an appropriate control, while Burke teaches that Cox5b mRNA increases and Cyc1 mRNA decreases in anaerobic conditions. Applicant therefore submits that Burke does not meet the limitations of claim 93 or claims dependent thereon, e.g., claims 106 and 107. Withdrawal of the rejection is respectfully requested.

CONCLUSION

Applicant respectfully requests entry of the amendment and consideration of the above remarks. In view of the above amendments and remarks, it is believed that all claims are allowable, and a notice to that effect is respectfully requested. If the Examiner feels that a telephone conference would expedite prosecution of this case, the Examiner is invited to call the undersigned at (617) 951-7000. Since an equal number of claims were canceled as were added, it is believed that no excess claims fees are due. A petition for a three-(3)-month extension of time is enclosed. Please charge any additional fees to our Deposit Account No. 18-1945, under Order No. WIBL-P01-013 from which the undersigned is authorized to draw.

Dated: December 9, 2008

Respectfully submitted,

By 

David P. Halstead, J.D., Ph.D.
Registration No.: 44,735

Customer Number: 28120

ROPES & GRAY LLP
One International Place
Boston, Massachusetts 02110-2624
(617) 854-2697
(617) 951-7050(Fax)
Attorneys/Agents For Applicant

RESOURCE

nature biotechnology

Large-scale chemical dissection of mitochondrial function

Bridget K Wagner[1,5], Toshimori Kitami[1,2,5], Tamara J Gilbert[1], David Peck[1], Arvind Ramanathan[1], Stuart L Schreiber[1], Todd R Golub[1,3] & Vamsi K Mootha[1,2,4]

Mitochondrial oxidative phosphorylation (OXPHOS) is under the control of both mitochondrial (mtDNA) and nuclear genomes and is central to energy homeostasis. To investigate how its function and regulation are integrated within cells, we systematically combined four cell-based assays of OXPHOS physiology with multiplexed measurements of nuclear and mtDNA gene expression across 2,490 small-molecule perturbations in cultured muscle. Mining the resulting compendium revealed, first, that protein synthesis inhibitors can decouple coordination of nuclear and mtDNA transcription; second, that a subset of HMG-CoA reductase inhibitors, combined with propranolol, can cause mitochondrial toxicity, yielding potential clues about the etiology of statin myopathy; and, third, that structurally diverse microtubule inhibitors stimulate OXPHOS transcription while suppressing reactive oxygen species, via a transcriptional mechanism involving PGC-1α and ERRα, and thus may be useful in treating age-associated degenerative disorders. Our screening compendium can be used as a discovery tool both for understanding mitochondrial biology and toxicity and for identifying novel therapeutics.

Oxidative phosphorylation (OXPHOS) is the core mitochondrial pathway responsible for ATP synthesis. The OXPHOS system consists of ~90 protein components, including all 13 of the proteins that are encoded by the mitochondrial genome (mtDNA)[1]. On time scales of seconds to minutes, mitochondrial ATP synthesis is regulated primarily by substrate availability and allosteric control[2]. During growth and development, however, transcription and translation of OXPHOS genes are carefully orchestrated between the nuclear and mitochondrial genomes to achieve sustained, metabolic adaptation. Over 50 mutations in the mtDNA and nuclear genome have been linked to rare, but devastating, inborn errors of OXPHOS metabolism[3]. Moreover, quantitative declines in OXPHOS activity and efficiency have been linked to nearly all age-associated degenerative diseases, including type 2 diabetes mellitus[4–6]. Hence, understanding OXPHOS function and regulation, particularly within the context of the entire cell, has important implications for managing many human diseases.

Traditional approaches to studying energy metabolism in the mitochondrion have focused either on the kinetics of ATP synthesis in isolated mitochondria or on transcriptional control of mitochondrial components. For instance, classic bioenergetic studies that assessed the effects of chemical inhibitors on isolated mitochondria[2] focused on the acute regulation of mitochondrial activity, ignoring the cell's ability to respond and adapt over longer time scales. Many of the chemical reagents used in these studies were incompatible with cellular or whole-animal studies, making it difficult to extend their relevance *in vivo*. More recently, molecular studies have shed light on the transcriptional control of OXPHOS[7], although relatively little is known about how these regulatory programs are coupled to other cellular processes.

Our goal was to combine physiologic and genomic profiling of intact cells in order to probe OXPHOS function and regulation in response to thousands of small-molecule perturbations. A systematic mapping between physiology and gene expression may shed light on how the nuclear genome and mtDNA are coordinated under different cellular conditions to maintain energy homeostasis. Large-scale approaches based on perturbing cells have been used successfully to study cell growth[8], gene expression[9,10] and tumorigenesis[11], but no previous studies appear to have combined cell-based, physiological readouts with genomic profiling. Our integrated approach provides a richer description of cellular mitochondrial state, reporting on more stable changes in the organelle, which can be useful for studying its longer-term adaptations.

Here, we report the construction of a mitochondrial screening compendium and its application to problems in mitochondrial biology, drug toxicity and therapeutics. Our compendium is freely available and should provide a framework for understanding how the activity and regulation of mitochondrial OXPHOS are integrated within the context of the entire cell.

RESULTS

Systematic profiling of OXPHOS physiology and transcription

We developed high-throughput assays in 384-well format for cell viability, mitochondrial physiology and gene expression (Fig. 1, Methods). We used differentiated C2C12 murine myotubes, a cell

[1]Broad Institute of Massachusetts Institute of Technology and Harvard, Seven Cambridge Center, Cambridge, Massachusetts 02142, USA. [2]Center for Human Genetic Research, Massachusetts General Hospital, 185 Cambridge Street, Boston, Massachusetts 02114, USA. [3]Department of Pediatric Oncology, Dana-Farber Cancer Institute, 44 Binney Street, Boston, Massachusetts 02115, USA. [4]Department of Systems Biology, Harvard Medical School, 200 Longwood Avenue, Boston, Massachusetts 02446, USA. [5]These authors contributed equally to this work. Correspondence should be addressed to V.K.M. (vamsi@hms.harvard.edu).

Received 26 November 2007; accepted 31 January 2008; published online 24 February 2008; corrected after print 8 July 2008; doi:10.1038/nbt1387

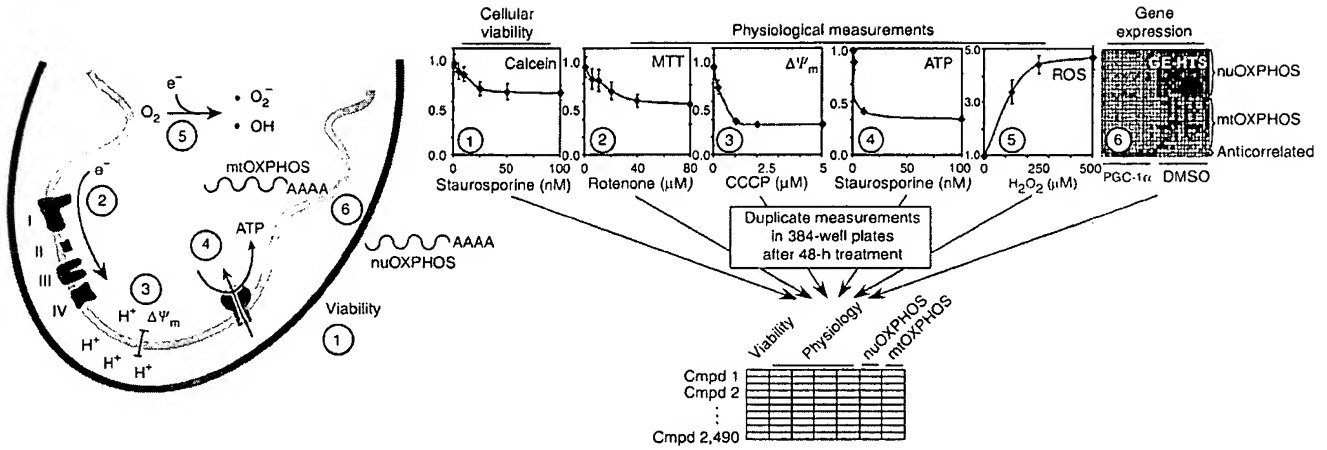


Figure 1 Complementary profiles of viability, mitochondrial physiology and gene expression across 2,490 chemical perturbations. The calcein assay (1) measures cell viability and filters out overtly toxic compounds, such as staurosporine. The MTT assay (2) measures cellular dehydrogenase activity, which is inhibited by the complex I inhibitor rotenone. The JC-1 assay (3) measures the mitochondrial membrane potential ($\Delta\Psi_m$) and drops acutely after the addition of the mitochondrial uncoupler carbonyl cyanide *m*-chlorophenylhydrazone (CCCP). A luciferase-based assay measures ATP (4), which is reduced by staurosporine. CM-H_2DCFDA is a fluorescent probe of cellular ROS (5), which can be stimulated by the addition of H_2O_2. The expression of both nuOXPHOS and mtOXPHOS transcripts is measured by a multiplex PCR technique, GE-HTS (6). Each column of the heat map represents one sample replicate; expression levels for each gene are row-normalized. Treatment with PGC-1α, an inducer of OXPHOS gene expression, is used as a positive control. All assays were performed in biological duplicate in 384-well format after 48 h of treatment in differentiated murine C2C12 myotubes. Data from 2,490 distinct compounds are incorporated into the screening compendium.

line that has been used extensively to study mitochondrial biogenesis and bioenergetics[12]. We measured cell viability using a calcein dye[13] and monitored mitochondrial physiology by probing four parameters related to OXPHOS. These included the JC-1 assay, which measures mitochondrial membrane potential ($\Delta\Psi_m$)[14]; the MTT ((3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide)) assay, which measures mitochondrial dehydrogenase activity[15]; a luciferase-based assay for cellular ATP levels[16]; and a fluorescent probe (CM-H_2DCFDA, (5-(and-6)-chloromethyl-2′,7′-dichlorodihydrofluorescein diacetate), which measures reactive oxygen species (ROS), a by-product of OXPHOS[17] (see Methods). We also developed an immunofluorescence-based assay for cytochrome *c* protein content (see Methods).

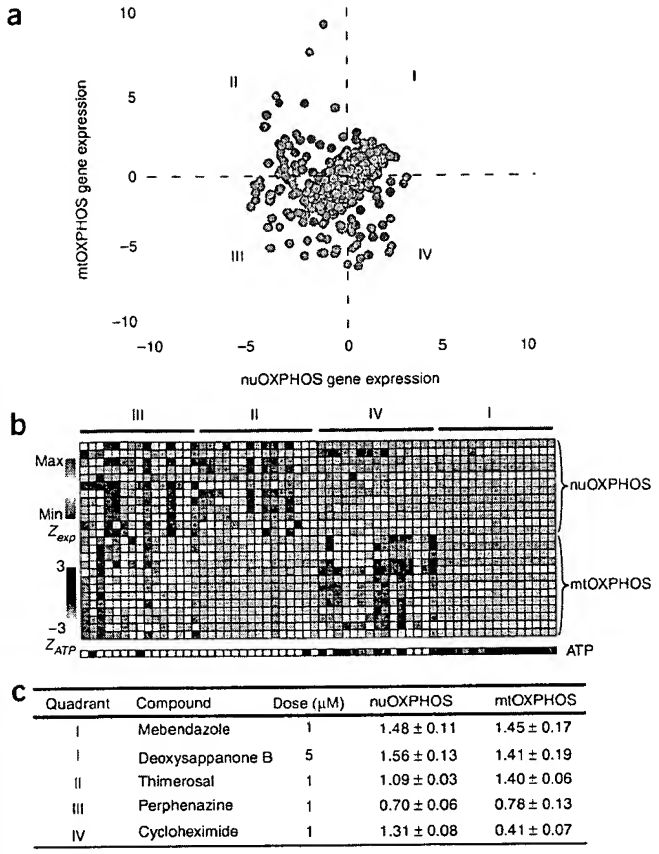
To complement these physiological assays, we also performed gene expression–based high-throughput screening (GE-HTS)[18,19] to profile transcripts associated with nuclear and mtDNA OXPHOS (see Methods). GE-HTS is a facile, high-throughput method that quantifies dozens of transcripts simultaneously. It is a multiplexed PCR strategy that combines ligation-mediated amplification with multicolored bead detection to identify and quantify transcripts of interest (see **Supplementary Fig. 1** online). We adapted GE-HTS to profile simultaneously all 13 mtDNA-encoded OXPHOS (mtOXPHOS) transcripts as well as 12 nuclear-encoded OXPHOS (nuOXPHOS) transcripts (**Supplementary Fig. 1**). These 12 nuOXPHOS transcripts include representatives from all five OXPHOS protein complexes and were selected because they capture virtually all of the variation in gene expression shown by the entire OXPHOS repertoire, as assessed by analysis of over 5,000 genome-wide microarrays (data not shown). Of note, our GE-HTS assay also monitored transcripts that tend to be anticorrelated to OXPHOS expression or are invariant across many conditions as assessed by microarray assays, and thereby assist in data analysis (see Methods). Together, our GE-HTS assay faithfully 'tags' the expression of the entire OXPHOS system. Moreover, because the expression of OXPHOS genes is so highly correlated, measuring multiple transcripts increases the signal-to-noise ratio with which we can detect subtle effects of individual compounds[4].

We performed the viability, physiology and gene-expression assays in duplicate in differentiated C2C12 myotubes following 48-h treatment with each of 2,490 compounds. Our chemical library consists of known bioactives, two-thirds of which are marketed drugs. Using a scoring algorithm dependent upon the distribution of mock-treated (DMSO) wells[20,21], we arrived at a normalized score for each assay in each well (**Supplementary Table 1** online). In our compendium, we have included data from our screen for cytochrome *c* protein expression, though we excluded it from subsequent analyses owing to the high coefficient of variation. Correlation analysis indicated that our remaining readouts (one for viability, four for OXPHOS physiology and one for OXPHOS gene expression) provide complementary information (**Supplementary Fig. 2** online).

The resulting compendium is freely available to the public and is useful for dissecting mitochondrial function (**Supplementary Table 1**). Unlike traditional approaches for studying mitochondrial function, it enables us to track systematically how changes in nuclear and mitochondrial OXPHOS gene expression are coupled to mitochondrial physiology over thousands of perturbations. We used this approach to explore three problems focused on mitochondrial biology, drug toxicity and the identification of novel therapeutics.

Cross-talk between nuclear and mitochondrial genomes

First, we used the compendium to identify the cellular signals involved in coordinating nuclear OXPHOS (nuOXPHOS) and mtDNA OXPHOS (mtOXPHOS) transcription. Expression of OXPHOS genes from the two genomes must be tightly coupled to maintain energy homeostasis in the mitochondrion[22,23]. Moreover, although OXPHOS expression can change in human diseases, it is often unclear whether the changes are primary or reactive and how these changes relate to cellular physiology[4,24–26]. We therefore focused on the relationship between nuOXPHOS and mtOXPHOS transcripts across



c

Quadrant	Compound	Dose (μM)	nuOXPHOS	mtOXPHOS
I	Mebendazole	1	1.48 ± 0.11	1.45 ± 0.17
I	Deoxysappanone B	5	1.56 ± 0.13	1.41 ± 0.19
II	Thimerosal	1	1.09 ± 0.03	1.40 ± 0.06
III	Perphenazine	1	0.70 ± 0.06	0.78 ± 0.13
IV	Cycloheximide	1	1.31 ± 0.08	0.41 ± 0.07

Figure 2 Coupling of nuclear and mitochondrial OXPHOS expression. (a) A two-dimensional plot of the composite *Z*-scores for nuOXPHOS and mtOXPHOS expression is shown. (b) Row-normalized heat map displaying the top 15 compounds in each quadrant (I–IV). Heat map of nuOXPHOS and mtOXPHOS expression is shown along with ATP levels. (c) Real-time PCR validation of select compounds at the indicated doses, using *Atp5a1* (nuOXPHOS) and *mt-Co1* (mtOXPHOS) normalized to *Hprt1* (internal control). Values indicate average fold change from mock-treated (DMSO) wells ± s.d. in four biological replicates.

the chemical perturbations. As expected, the majority of compounds influence the two sets of genes in a coordinated manner (**Fig. 2a**). However, we identified some compounds that decouple the coordination between these two genomes (**Fig. 2b** and **Supplementary Table 2** online), a subset of which we confirmed with follow-up dose response curves and RT-PCR analysis (**Fig. 2c**). Specifically, we discovered that the eukaryotic protein synthesis inhibitors emetine, anisomycin and cycloheximide preferentially increase nuOXPHOS expression, implying that translational control might be important in coordinating the two genomes. Follow-up studies revealed that 1 μM cycloheximide elevated nuOXPHOS 1.3-fold but decreased mtOXPHOS 2.4-fold (**Fig. 2c**). Notably, we found that nuOXPHOS expression, but not mtOXPHOS expression, correlated strongly with cellular ATP levels (**Fig. 2b**). To determine whether nuOXPHOS expression drives the changes in ATP levels, or reacts to changes in ATP levels, we performed follow-up time-course analyses with 20 μM perphenazine, a compound that decreased nuOXPHOS expression. Whereas nuOXPHOS expression declined significantly (21%, *t*-test, $P = 0.004$) within the first hour of treatment, cellular ATP levels remained unchanged (0.6%, *t*-test, $P = 0.84$) at these early time points. At later time points, however, ATP levels dropped significantly (8 h: 11% decrease, *t*-test, $P = 1.4 \times 10^{-5}$, 24 h: 27% decrease, *t*-test, $P = 6.3 \times 10^{-22}$), suggesting that the decline in nuOXPHOS expression precedes and drives the decline in cellular ATP levels.

Exploring the mitochondrial basis for drug toxicity

To probe the role of mitochondria in human drug toxicity, we focused on the statins—3-hydroxy-3-methylglutaryl–coenzyme A (HMG-CoA) reductase inhibitors taken by nearly 100 million patients worldwide. Statins are associated with a 0.1–0.5% incidence of myopathy[27], believed to be caused by ubiquinone depletion, which can block electron transport[28]. However, clinical and epidemiological studies of the association between statins and myopathy have produced conflicting results[28]. Of the six statins present in our screening collection, three (fluvastatin, lovastatin and simvastatin) produced strong decreases in cellular ATP levels and MTT activity (**Fig. 3a**). Previous studies showed that lovastatin and simvastatin reduce MTT activity and ATP levels[29], consistent with our high-throughput screening results. To eliminate the possibility that we uncovered two classes based merely on potency, we measured cellular ATP levels over doses ranging up to 40 μM. We observed the same segregation of effects, with atorvastatin, pravastatin and rosuvastatin showing little to no effect on cellular or mitochondrial ATP levels (**Supplementary Fig. 3** online).

To determine whether this profile might represent a signature of drug-induced myopathy, we established a centroid profile for the three mitochondria-active statins (fluvastatin, lovastatin and simvastatin) and sought to identify other clinically used drugs with a similar assay profile. The ten nearest-neighbor drugs to the centroid statin profile (**Fig. 3b**) were amoxapine, cyclobenzaprine, propranolol, griseofulvin, pentamidine, paclitaxel, propafenone, ethaverine, trimeprazine and amitriptyline. Notably, five of these compounds (amoxapine, propranolol, griseofulvin, pentamidine and paclitaxel) have also been associated with skeletal muscle myopathy or myalgia[30–34], a strikingly high proportion in comparison to the small fraction of all FDA-approved drugs believed to be associated with this side effect[35]. This suggests that the drug profile might be indicative of myopathy or myalgia. Further examination of the screening data revealed that two electron transport chain inhibitors—β-dihydrorotenone (a complex I inhibitor) and antimycin A (a complex III inhibitor)—were among the 16 nearest-neighbor compounds to this assay profile, which provides mechanistic insight into this profile. Together, the data support the idea that myopathy induced by these five other drugs could be mitochondrial in origin.

We were struck by the fact that one of these nearest-neighbor drugs is propranolol, a widely used antihypertensive agent. Follow-up experiments confirmed that propranolol, but not other selective β_1 blockers, decreases cellular ATP levels in a dose-dependent manner (**Supplementary Fig. 3**). Because many patients take both a statin and a β-blocker for cardioprotection, we were curious whether the two drugs might interact to cause toxicity. We thus assessed cellular ATP levels after treatment with all possible combinations of the six statins in our collection and three β-blockers (atenolol, metoprolol and propranolol), with all concentrations falling between 2.5 and 10 μM (**Fig. 3c**). Although neither atenolol nor metoprolol showed an effect either alone or in combination with any statin, propranolol had an additive effect on statin-induced decrease in ATP levels, as determined using the Bliss independence model (**Fig. 3c**)[36]. Our screening compendium and follow-up experiments (**Fig. 3c**) thus raise the potentially important hypothesis that patients on a combination of propranolol and one of the three statins (fluvastatin, lovastatin, simvastatin) might be at a higher risk for developing myopathy or myalgia.

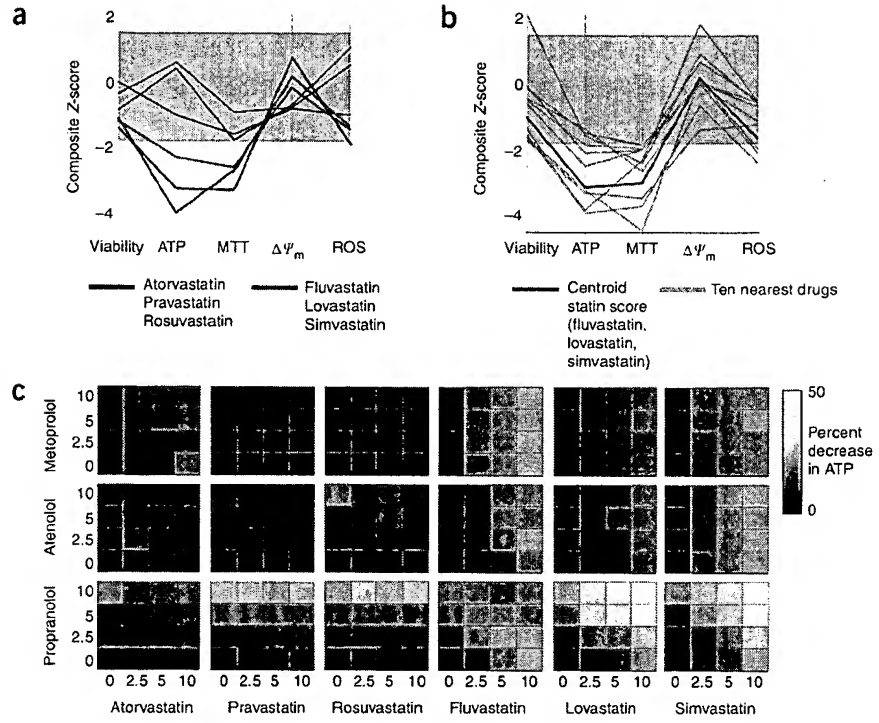


Figure 3 Statin-induced mitochondrial toxicity. (**a**) Six of the HMG-CoA reductase inhibitors (statins) in clinical use are in the chemical screening collection. Composite *Z*-scores for cell viability, ATP generation, MTT activity, $\Delta\Psi_m$, ROS levels and gene expression are shown, where negative scores indicate a decrease in signal compared to mock-treated (DMSO) wells. The gray shading indicates scores that fall within the noise envelope. (**b**) A centroid statin score was generated by calculating the arithmetic means of the composite *Z*-scores for fluvastatin, lovastatin and simvastatin. The ten nearest-neighbor clinically used drugs (amoxapine, cyclobenzaprine, propranolol, griseofulvin, pentamidine, paclitaxel, propafenone, ethaverine, trimeprazine and amitriptyline) were identified by calculating the root-mean-square distance of each performance vector to the profile of interest. (**c**) All six statins were tested in combination with three clinically used β-adrenergic blockers (propranolol, atenolol and metoprolol) for their effects on cellular ATP levels. Compound concentrations are indicated on each axis, and the color indicates the change in ATP levels (ranging from black, for no change, to yellow, for a 50% decrease). Data represent the average of six independent replicates; coefficients of variation were all below 15%.

Discovering lead compounds for common degenerative diseases

Finally, we searched for lead compounds that might hold therapeutic potential for common human diseases. As we and others have recently shown that a decline in OXPHOS gene expression and an elevation in ROS generation are associated with type 2 diabetes[4,37,38], neurodegeneration[39] and aging[5,6], we queried our compendium to identify compounds that might be capable of elevating OXPHOS expression while reducing ROS accumulation.

We used two computational strategies to spotlight such compounds (see Methods). First, we developed a simple analytical strategy to determine whether any structurally related set of compounds might boost OXPHOS expression while also suppressing ROS accumulation (**Supplementary Table 3** online). This strategy involves organizing all compounds based on structural similarity (Methods) and then asking whether members of a cluster had concordant scores in a given assay (**Fig. 4a**). The advantage of this strategy is that individual compounds might show a subtle response not detectable in a primary screen with duplicate measurements, whereas the grouped analysis provides added statistical power. Second, in a complementary approach, we sought to identify individual compounds that promote OXPHOS gene

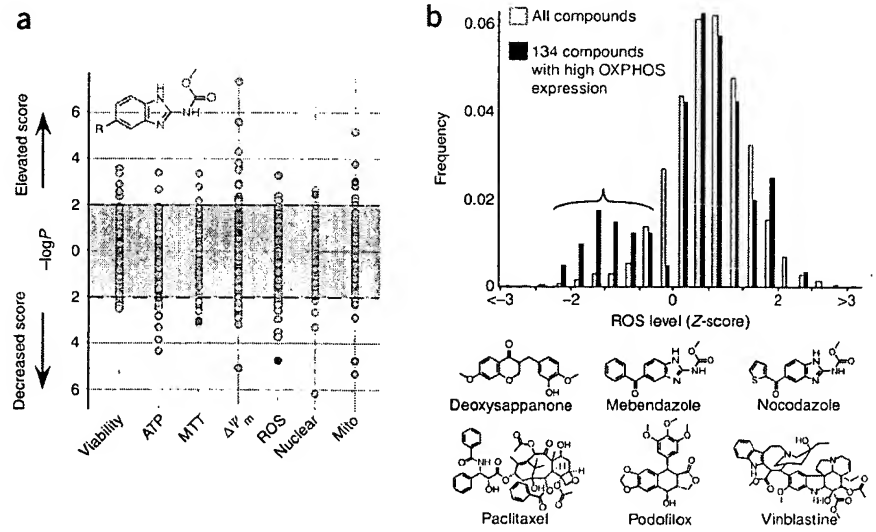


Figure 4 Two complementary strategies to identify small molecules that boost OXPHOS gene expression and decrease ROS levels. (**a**) Mining the compendium for sets of structurally related compounds that achieve the desired activity. All compounds were organized into 624 clusters based on the chemical descriptors molecular weight, log *P*, number of hydrogen bond donors and acceptors, and number of rotatable bonds. The Mann-Whitney rank-sum statistic for each cluster and each assay was then calculated. The significance of each cluster in each assay is shown, with points above zero indicating positive composite scores and points below zero showing negative composite scores. A nominal $P = 0.01$ is delimited by the dashed lines. The red data points spotlight a single cluster that is significant for the desired activity, with the shared chemical scaffold shown. (**b**) Mining the compendium for individual compounds that achieve the desired activity. The distributions of ROS scores are shown for all compounds (gray) and for compounds associated with the highest OXPHOS gene expression (black). The latter follow a bimodal distribution, and the smaller mode (bracketed) contains six compounds that elevate OXPHOS expression and decrease ROS levels, with chemical structures shown.

expression while reducing ROS levels. The advantage of this method is that it can reveal structurally unrelated compounds that individually exert large effect sizes in the two assays of interest. We focused on the upper tail of the combined nuclear and mitochondrial OXPHOS gene-expression distribution (**Fig. 4b**).

Notably, both analytical strategies spotlighted microtubule modulators, including both a microtubule stabilizer (paclitaxel) and several destabilizers (mebendazole, nocodazole, podophyllotoxin and vinblastine), as agents that boost OXPHOS expression while suppressing ROS levels. The second strategy also yielded deoxysappanone B, a natural product found in sappan wood[40], whose molecular mode of action is unknown and has not been previously linked to microtubule biology. The other microtubule inhibitors within the compound collection (colchicine and griseofulvin) did not display the same decrease in ROS levels, but did show a modest increase in OXPHOS expression.

Next, we were interested in confirming these primary screening results and determining whether the effects on OXPHOS expression and ROS levels occur via shared or distinct mechanisms, and whether these were on-target or off-target effects of microtubule disruption. We therefore retested the microtubule modulators at a range of 20 nM to 20 μM (**Fig. 5a**). Treatment with either deoxysappanone B, mebendazole, nocodazole, podophyllotoxin or vinblastine increased OXPHOS expression and decreased ROS levels at the same dose of 2 μM. In contrast, paclitaxel showed effects in the two assays at 20 nM, suggesting a shared mechanism for OXPHOS expression and ROS level. Notably, at these doses, these compounds did not decrease cell viability (**Fig. 5a**), indicating that the decline in ROS is not simply a reflection of overt cytotoxicity. We also imaged tubulin immunofluorescence after treatment with deoxysappanone B and paclitaxel, two compounds that showed distinct potencies. For both compounds, the potency required for microtubule disruption was the same as that

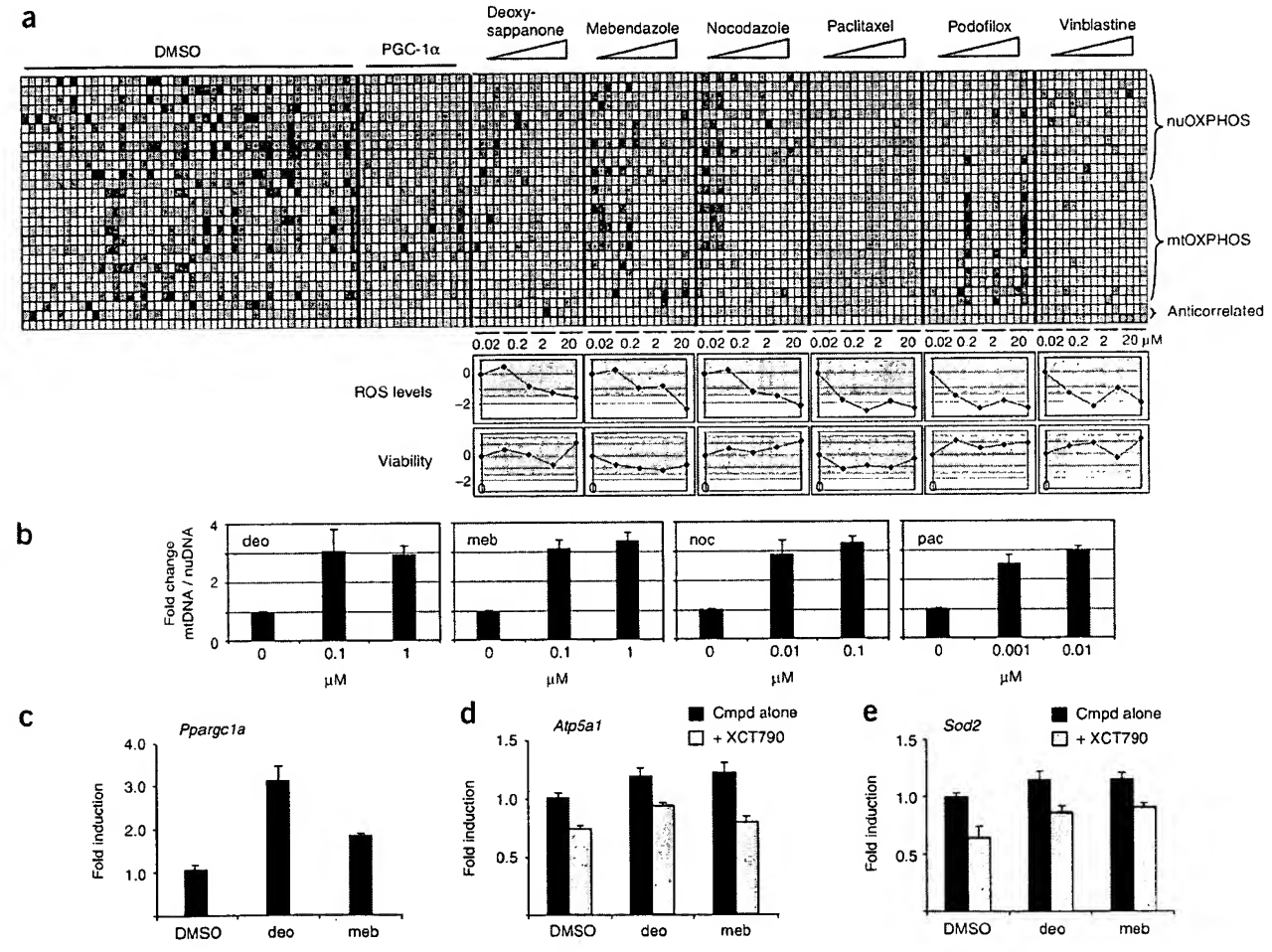


Figure 5 Secondary analyses of the effects of microtubule inhibitors on OXPHOS gene expression and physiology. (**a**) Compounds indicated in Figure 4b were retested at 20 nM, 200 nM, 2 μM and 20 μM. Gene expression levels are represented as a row-normalized heat map, with negative controls (DMSO treatment) and positive controls (PGC-1α treatment) shown. Dose-response curves for ROS levels and viability are also provided, where the *y*-axis is the composite *Z*-score. Shaded area indicates the noise envelope ($P > 0.05$). Data shown are the results of four biological replicates per concentration. (**b**) Analysis of mtDNA/nuDNA copy number ratio after treatment with four of the compounds (deo, deoxysappanone B; meb, mebendazole; noc, nocodazole; pac, paclitaxel), using three biological replicates, normalized to DMSO treatment alone. (**c**) Quantitative PCR measurement of *Ppargc1a* gene expression, in response to either DMSO alone, 5 μM deoxysappanone B (deo) or 1 μM mebendazole (meb). (**d**) Quantitative PCR measurement of the nuclear OXPHOS gene *Atp5a1*. Cells were either treated with compound alone (black bars) or in combination with 5 μM of the ERRα inverse agonist XCT790 (gray bars). (**e**) Quantitative PCR measurement of *Sod2*, which encodes the ROS scavenger MnSOD, as in **d**. Means and s.d. of expression data are the result of four biological replicates.

required to affect OXPHOS expression and ROS levels (**Supplementary Fig. 4** online). To our knowledge, deoxysappanone B has not previously been linked to microtubule inhibition, but it now has been predicted to do so and the prediction has been validated by this study. Given that structurally and mechanistically diverse microtubule modulators increased OXPHOS gene expression, decreased cellular ROS and disrupted microtubules with equivalent potencies, it is likely that these effects are directly related to inhibition of microtubules, and not due to an off-target effect.

Because mtDNA replication and transcription are often coupled[41], we sought to determine whether any of these compounds promoted mtDNA replication. At the concentrations tested, several of these microtubule modulators—but not podophyllotoxin or vinblastine—increased mtDNA copy number approximately threefold (**Fig. 5b**).

We sought to determine the transcriptional mechanism by which microtubule inhibition might promote OXPHOS expression and mtDNA replication while suppressing ROS. We hypothesized that these changes might be occurring via PGC-1α, a transcriptional coactivator that regulates mitochondrial biogenesis in muscle[42] and whose transcriptional program is diminished in type 2 diabetes[4]. Consistent with this hypothesis, both mebendazole and deoxysappanone B induced the expression of *Ppargc1a* (which encodes PGC-1α) by approximately threefold (**Fig. 5c**). We have previously shown that the transcription factor ERRα serves as a key transcriptional partner of PGC-1α to drive OXPHOS expression in muscle, and that disruption of ERRα with the selective inverse agonist XCT790 suppresses PGC-1α–induced OXPHOS expression[43]. Therefore, we tested whether XCT790 is capable of inhibiting compound-induced transcription. We observed that both mebendazole and deoxysappanone B increased the expression of a nuclear OXPHOS gene, *Atp5a1*, by ~20%, and that the increase by mebendazole was blunted by XCT790 (**Fig. 5d**), further suggesting a PGC-1α–dependent mechanism of compound activity. The mitochondrial ROS scavenger MnSOD is downstream of the same PGC-1α pathway[44], and as we observed decreased cellular ROS levels after treatment with these small molecules, we also tested the effects of the compounds on this gene. A similar increase in MnSOD levels, which was only partially suppressible by XCT790, was observed with these compounds (**Fig. 5e**). These results suggest that microtubule modulators both activate OXPHOS transcription and reduce cellular ROS levels in a manner involving PGC-1α and ERRα.

DISCUSSION

The mitochondrion is an extremely complex organelle, with components derived from both the nuclear and mitochondrial genomes, whose activity must be carefully coupled to cellular metabolism and signaling. We systematically investigated mitochondrial function using multiple physiological and multiplexed gene-expression assays for OXPHOS following chemical perturbations. The gene-expression and physiological assays provided complementary information and were jointly useful in interpreting the effects of compound treatment. The compendium is freely available and can be used to investigate the network properties of the mitochondrion. Moreover, because so many of the 2,490 compounds tested are well characterized, they can be used immediately to study mitochondrial biology, as illustrated by the three applications we have demonstrated.

First, we focused on how the nuclear genome and mtDNA are coordinated across a variety of physiological states. As this is the first compendium to interrogate the expression of both the nuclear genome and mtDNA, we could show that the bulk of compounds coordinately regulate expression from both genomes. However, similar to the demonstration that the calcium ionophore A-23187 can elevate nuOXPHOS while decreasing mtOXPHOS[45], we found that eukaryotic protein synthesis inhibitors disrupt cross-talk between these two genomes. We now have an array of chemical tools (Fig. 2) to investigate whether protein synthesis inhibitors also disrupt the nuclear-to-mitochondrial genome cross-talk via known pathways or through one or more novel mechanisms.

Second, we mined the compendium to learn about drug toxicity. Our focus on statins revealed a bimodal response from these HMG-CoA reductase inhibitors (**Fig. 3**), with three compounds (atorvastatin, pravastatin and rosuvastatin) forming a group with no activity in our physiological assays and three other compounds (fluvastatin, lovastatin and simvastatin) forming a group showing signs of OXPHOS inhibition. Statins block the synthesis of mevalonate, a precursor not only of cholesterol but also ubiquinone, a mobile electron carrier in the mitochondrion that is critical to OXPHOS function. It has therefore been previously suggested that statin myopathy may result from a respiratory chain blockade[28]. Our data suggest that the statins fluvastatin, lovastatin and simvastatin should be examined further for their effects on skeletal muscle ubiquinone and drug-induced myopathy. Using the statin toxicity signature (**Fig. 3a**), we also discovered that propranolol, a widely used antihypertensive agent, shares the same profile of mitochondrial inhibition. β_1-selective blockers such as atenolol and metoprolol, however, did not show this pattern of toxicity (**Supplementary Fig. 3**). The additive interaction we reveal between the statins and propranolol suggests that patients taking both statins and propranolol might be at increased risk for developing skeletal muscle myopathy or myalgia. Because many patients with heart disease are likely to be on this drug combination, our hypothesis can be tested easily and may help to account for the conflicting reports on skeletal muscle myopathy associated with statins.

Third, we queried our compendium for compounds that could potentially reverse signatures associated with age-associated degenerative diseases. We and others have recently reported that a decline in OXPHOS expression and a rise in ROS levels accompany a number of common degenerative diseases, including diabetes, neurodegeneration and aging[4–6,37–39]. Using two computational approaches to mine our compendium, we discovered that structurally unrelated microtubule modulators with diverse mechanisms of action (both stabilizers and destabilizers) increased OXPHOS expression and decreased cellular ROS. Follow-up dose-response studies of mitochondrial function and microtubule disruption suggest that these drugs affect OXPHOS expression and ROS levels via an on-target mechanism.

Our studies raise the possibility that manipulation of the microtubule pathway may reverse the gene-expression and ROS signatures associated with common degenerative diseases and that these may represent therapeutic targets. Previous clinical case reports support this hypothesis. Type 1 and type 2 diabetics treated for parasitic worm infections with mebendazole unexpectedly showed improvement both in fasting blood glucose levels and nonesterified fatty acid metabolism[46]. Blood glucose levels improved without weight loss, suggesting that the improvement in diabetes was not due to a toxic side effect of mebendazole. Moreover, deoxysappanone B, a natural product found in sappan wood, has been reported to be an antidiabetic component of Chinese herbal medicine, although its mechanism was unclear[40]. Our data convincingly demonstrate that deoxysappanone B disrupts microtubules (**Supplementary Fig. 4**).

At a molecular level, we have uncovered an unexpected link between microtubule disruption and an increase in PGC-1α/ERRα–mediated OXPHOS gene expression. Although changes in mitochondrial staining and morphology have been associated with microtubule inhibitors[47], no studies have specifically documented their effects on

OXPHOS expression and ROS levels. It is intriguing to speculate that perhaps interactions between the cytoskeleton and the mitochondrion are important in integrating cellular homeostasis throughout the cell cycle. As many of these microtubule modulators are used for treating cancer, our results may enhance understanding of the metabolic basis of chemotherapeutic action.

Building on decades of research in bioenergetics that yielded a detailed, mechanistic understanding of OXPHOS in isolated mitochondria[48], our screening compendium provides a foundation for understanding how OXPHOS physiology and regulation are integrated within the broader signaling and metabolic network of the cell. Each small-molecule probe modulates a different aspect of cell biology, and our compendium provides insights into how OXPHOS responds to these perturbations by remodeling at the levels of physiology and gene expression. Because many of the small molecules used in this study are well-characterized bioactives, our compendium provides a rich chemical toolkit for manipulating mitochondria in cells and whole animals in a defined manner. Given the growing number of rare and common diseases associated with different types of mitochondrial dysfunction, we anticipate that our compendium will serve as a generic tool for systematically investigating mitochondrial pathogenesis and for restoring mitochondrial function.

METHODS

Cell culture. C2C12 myoblasts (ATCC) were grown in Dulbecco's Modified Eagle's Medium (DMEM, Mediatech) supplemented with 10% (vol/vol) FBS and antibiotics (100 μg/ml penicillin/streptomycin mix) in a humidified atmosphere at 37 °C with 5% CO_2. Differentiation into myotubes was induced at 80% density on 'day 0' by changing the medium to DMEM supplemented with 2% (vol/vol) horse serum.

Cell-based high-throughput screening. For all screening, 4,000 C2C12 myoblasts per well were seeded into either black or white 384-well optical-bottom plates (Nunc) at 50 μl per well. On day 4 of differentiation, 100 nl of each compound was pin-transferred in duplicate into fresh medium with a steel pin array, using the CyBi-Well robot (CyBio). To increase the number of mock-treated wells included in the control distribution, we added an additional plate containing DMSO alone. Compound-treated plates were incubated at 37 °C for 48 h. All cell-based assay measurements were performed using the EnVision plate reader (PerkinElmer). The coefficient of variation for each of these assays was estimated to be less than 15%. All data has been deposited in ChemBank: http://chembank.broad.harvard.edu/assays/viewproject.htm?id=1000453.

Calcein viability assay. Medium was aspirated from plates, and 30 μl per well 1 μM calcein-AM (Molecular Probes) in phenol red–free medium was added. Plates were incubated for 1 h at 37 °C and washed three times with 50 μl per well PBS. Fluorescence was measured at excitation and emission wavelengths (ex/em) of 485 nm/530 nm.

JC-1 mitochondrial membrane potential assay. Upon depolarization, the JC-1 dye is converted from a diffuse green form to red fluorescent J-aggregates. The ratio of red to green fluorescence serves as a readout of the mitochondrial membrane potential. Medium was aspirated from plates, and 20 μl per well 3.25 μM JC-1 (Molecular Probes) in phenol red–free medium was added. Plates were incubated for 2 h at 37 °C and washed three times with 50 μl per well PBS. Fluorescence was measured first at ex/em 530 nm/580 nm ('red') and then at ex/em 485 nm/530 nm ('green').

Assay for cellular ATP levels. 20 μl per well CellTiterGlo reagent (Promega) was added to 20 μl per well of cell culture medium. Plates were agitated for 2 min and incubated for 10 min at room temperature (22–24 °C) before luminescence was measured.

MTT assay. Medium was aspirated from plates, and 50 μl per well 0.5 mg/ml MTT in phenol red–free medium was added. Plates were incubated for 2 h at 37 °C, and this was followed by aspiration of MTT solution, addition of 50 μl per well DMSO to dissolve formazan crystals, and incubation at 37 °C for 30 min. After incubation, plates were equilibrated to room temperature for an additional 20–30 min. Absorbance was measured at 540 nm.

Reactive oxygen species assay. Medium was aspirated from plates, and 20 μl per well 10 μM CM-H_2DCFDA (Molecular Probes) in phenol red–free medium was added. Plates were incubated for 1 h at 37 °C and washed three times with 50 μl per well PBS. Fluorescence was measured at ex/em 485 nm/530 nm.

Cytochrome c protein detection. Cells were fixed with 3.7% (vol/vol) formaldehyde in PBS for 30 min and then washed with TBS containing 0.1% (vol/vol) Tween-20 (TBST) and blocked with TBST + 3% (wt/vol) BSA for 1 h at room temperature. Cytochrome *c* was detected by incubating the cells with primary antibody (Cell Signaling Technology; 1:100) overnight at 4 °C, washing three times with TBST, and incubating with secondary antibody (Alexa Fluor 488–conjugated anti-mouse IgG, Invitrogen; 1:250) for 1 h at room temperature. Plates were washed three times with TBST and fluorescence measured at ex/em 485 nm/530 nm.

Gene expression–based high-throughput screening. We adapted the GE-HTS assay to monitor both nuclear and mtDNA OXPHOS transcripts. To narrow down the list of potential genes from nearly 80 nuclear OXPHOS genes, we used a list of highly co-regulated OXPHOS genes[4] that are coordinately expressed across tissues and are downstream of the PGC-1α transcriptional coactivator. From this list, we selected genes that showed the highest signal-to-noise ratio in the microarray analysis of PGC-1α overexpression in C2C12 myotubes[43] representing all five OXPHOS complexes. We also selected two genes that are downregulated by PGC-1α with the best signal-to-noise ratio. As controls, we selected genes that showed the lowest signal (no treatment effect) and lowest noise (biological variation) in the PGC-1α overexpression data[43], as well as genes previously found to be invariant from the analysis of multiple microarray datasets[49]. We selected control genes that span a wide range of expression levels to prevent biasing for abundant transcripts. The selected OXPHOS transcripts capture the bulk of the variation exhibited by the OXPHOS transcripts represented on over 5,000 publicly available mouse microarrays on the Affymetrix platform (data not shown).

From the list of OXPHOS genes and control genes for GE-HTS, we designed probe pairs with T7 and T3 universal primer sites, 40-bp target sequence split into two 20-bp sequences for each probe, and gene-specific barcode sequence attached to the 5′ probe according to the published assay specification[19]. We selected 40-bp gene-specific target sequences that are not alternatively spliced using oligonucleotide sequences found in the Mouse Exonic Evidence-Based Oligonucleotide Chip (MEEBO, http://alizadehlab.stanford.edu/). Full primer sequences are included in **Supplementary Table 4** online.

The GE-HTS assay was performed as previously described[50]. Because this assay measures the final amount of PCR products rather than providing a real-time measurement of gene expression, we adjusted the parameters in the original protocol so that the abundance of PCR products were within the linear range of the assay. We removed 20 μl of medium and added 25 μl of lysis buffer per well of a 384-well plate, and used 24 PCR cycles instead of the 29 cycles described[50]. We used 32 DMSO-treated and 32 PGC-1α adenovirus–treated wells per 384-well compound plate, with one additional control plate containing 192 DMSO-treated wells, 32 GFP adenovirus–treated wells and 160 PGC-1α adenovirus–treated wells. The PGC-1α adenovirus–treated cells serve as a positive control for increased OXPHOS gene expression, as previously reported[4].

Tubulin immunofluorescence. On day 4 of differentiation, C2C12 myotubes were treated with each compound for 48 h and then fixed for 5 min in ice-cold 100% methanol. Cells were washed once in 50 μl PBSTB2 (PBS with 0.1% (vol/vol) Tween-20 and 2% (wt/vol) BSA) and blocked in PBSTB2 for 1 h at room temperature or overnight at 4 °C. Cells were incubated with an anti-α-tubulin (Sigma-Aldrich) antibody, 1:1,000 in PBSTB2, for 1 h at room temperature, and then washed three times with PBSTB2. Cells were incubated with secondary antibody (Alexa 488–conjugated anti-mouse antibody, 1:500 in PBSTB2) (Molecular Probes) and Hoechst 33342 for 1 h at room temperature and then washed three times in PBSTB2. Cells were visualized using an automated microscope (IX-Micro, Molecular Devices).

Quantitative PCR of mtDNA and transcripts. *mtDNA quantification.* Mitochondrial DNA copy number was assessed by quantifying the abundance of the

mitochondrial gene *mt-Co1* (encoding cytochrome *c* oxidase 1) relative to the nuclear gene *Actb* (encoding β-actin). DNA from cells were extracted using DNeasy (Qiagen) and quantified for *mt-Co1* and *Actb* copy number using quantitative PCR (Applied Biosystems). The change in the *mt-Co1/Actb* ratio between the compound-treated and DMSO control cells represents the fold change in mtDNA copy number.

Gene expression. We extracted RNA using an RNeasy kit (Qiagen) and synthesized cDNA using a high-capacity cDNA reverse transcription kit (Applied Biosystems) with random hexamers, as described by the manufacturer. The cDNA was then used for real-time PCR quantification of products for mouse *Atp5a1* (Mm00431960_m1), *Sod2* (*MnSOD*; Mm01313000_m1) and *Ppargc1a* (Mm00447183_m1), with *Hprt1* (Mm03024075_m1) serving as an internal control, using TaqMan gene-expression assays (Applied Biosystems).

Statistics. *Cell-based screening.* Composite *Z*-scores reflecting compound performance as compared to a mock-treated (DMSO) distribution were calculated as described[20,21] (see also http://chembank.broad.harvard.edu/details.htm?tag=Help#screeningData).

GE-HTS. We first eliminated wells that failed the assay reaction by filtering out wells in which the raw expression value of *Rps2* (a control gene) was 2 s.d. below the median DMSO control value for each plate. We normalized for plate-to-plate variation by scaling the per-well expression level of each gene to the median expression level of that gene in PGC-1α control wells on each plate. We set the median PGC-1α–treated expression value for each gene to 1, and then normalized for well-to-well variation by dividing the expression level of each OXPHOS gene by the average value of eight control genes for each well. This number represents the processed data value.

To score the expression levels of 12 nuclear- and 13 mitochondrial-encoded OXPHOS genes, we first weighted each gene by its ability to distinguish DMSO control wells from PGC-1α-treated wells. We calculated the signal-to-noise ratio[49] of each gene using our PGC-1α-treated positive control and DMSO negative control, and multiplied the expression value of each gene per well by this signal-to-noise ratio. We then summed these weighted scores over nuclear-encoded or mitochondrial-encoded OXPHOS genes to derive one score each for expression within each genome. Composite *Z*-scores were calculated as described above.

Similarity between assay profiles. We used the cell-based composite *Z*-scores from the ATP, MTT, JC-1 and ROS assays to calculate the root-mean-square distance between performance vectors, as this statistic gives greater weight to values far from zero. We obtained centroid statin scores by taking the arithmetic mean of the composite *Z*-scores from these four assays.

Identifying structurally related small molecules. We used Pipeline Pilot (Scitegic) to perform *K*-means clustering of the molecules based on common and biologically intuitive chemical features (molecular weight, octanol-water partition coefficient, number of hydrogen bond donors and acceptors, and number of rotatable bonds). We set *K* to 624 to result in an average of 5 compounds per cluster. To detect enrichment for assay performance within each compound cluster, we performed the Mann-Whitney rank-sum test on each cluster in each assay.

Note: Supplementary information is available on the Nature Biotechnology website.

ACKNOWLEDGMENTS

We thank Stephanie Norton, Jason Burbank, Mariah Eustice and Nicky Tolliday for assistance in high-throughput screening; Nathan Billings and Olga Goldberger for technical assistance; Oded Shaham, Ken Ross and Paul Clemons for computational assistance; and Joel Hirschhorn, Eric Lander and Robert Gould for thoughtful discussions and comments on the manuscript. S.L.S. and T.R.G. are Investigators of the Howard Hughes Medical Institute. V.K.M. is recipient of a Career Award in the Biomedical Sciences from the Burroughs Wellcome Fund, a Charles E. Culpeper Scholarship in Medical Science, and a Physician Scientist Early Career Award from the Howard Hughes Medical Institute. This work was supported by grants from the National Institute of Health (National Institute of Diabetes and Digestive and Kidney Diseases), the American Diabetes Association and the Richard and Susan Smith Family Foundation (V.K.M.).

AUTHOR CONTRIBUTIONS

V.K.M. conceived of and supervised the project. B.K.W., T.K. and V.K.M. designed the experiments and analyzed the data. T.J.G., A.R. and B.K.W. carried out phenotypic screening. D.P. and T.K. carried out GE-HTS experiments. S.L.S. and T.R.G. provided guidance on chemical screening and GE-HTS, respectively, and advised on analysis. B.K.W., T.K. and V.K.M. wrote the paper.

Published online at http://www.nature.com/naturebiotechnology/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions

1. Anderson, S. *et al.* Sequence and organization of the human mitochondrial genome. *Nature* **290**, 457–465 (1981).
2. Chance, B. & Williams, G.R. Respiratory enzymes in oxidative phosphorylation. III. The steady state. *J. Biol. Chem.* **217**, 409–427 (1955).
3. DiMauro, S. & Schon, E.A. Mitochondrial respiratory-chain diseases. *N. Engl. J. Med.* **348**, 2656–2668 (2003).
4. Mootha, V.K. *et al.* PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* **34**, 267–273 (2003).
5. Petersen, K.F. *et al.* Mitochondrial dysfunction in the elderly: possible role in insulin resistance. *Science* **300**, 1140–1142 (2003).
6. Balaban, R.S., Nemoto, S. & Finkel, T. Mitochondria, oxidants, and aging. *Cell* **120**, 483–495 (2005).
7. Kelly, D.P. & Scarpulla, R.C. Transcriptional regulatory circuits controlling mitochondrial biogenesis and function. *Genes Dev.* **18**, 357–368 (2004).
8. Weinstein, J.N. *et al.* An information-intensive approach to the molecular pharmacology of cancer. *Science* **275**, 343–349 (1997).
9. Hughes, T.R. *et al.* Functional discovery via a compendium of expression profiles. *Cell* **102**, 109–126 (2000).
10. Lamb, J. *et al.* The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* **313**, 1929–1935 (2006).
11. Ramanathan, A., Wang, C. & Schreiber, S.L. Perturbational profiling of a cell-line model of tumorigenesis by using metabolic measurements. *Proc. Natl. Acad. Sci. USA* **102**, 5992–5997 (2005).
12. Leary, S.C., Battersby, B.J., Hansford, R.G. & Moyes, C.D. Interactions between bioenergetics and mitochondrial biogenesis. *Biochim. Biophys. Acta* **1365**, 522–530 (1998).
13. Dolma, S., Lessnick, S.L., Hahn, W.C. & Stockwell, B.R. Identification of genotype-selective antitumor agents using synthetic lethal chemical screening in engineered human tumor cells. *Cancer Cell* **3**, 285–296 (2003).
14. Smiley, S.T. *et al.* Intracellular heterogeneity in mitochondrial membrane potentials revealed by a J-aggregate-forming lipophilic cation JC-1. *Proc. Natl. Acad. Sci. USA* **88**, 3671–3675 (1991).
15. Berridge, M.V. & Tan, A.S. Characterization of the cellular reduction of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT): subcellular localization, substrate dependence, and involvement of mitochondrial electron transport in MTT reduction. *Arch. Biochem. Biophys.* **303**, 474–482 (1993).
16. Crouch, S.P., Kozlowski, R., Slater, K.J. & Fletcher, J. The use of ATP bioluminescence as a measure of cell proliferation and cytotoxicity. *J. Immunol. Methods* **160**, 81–88 (1993).
17. Ye, G., Metreveli, N.S., Ren, J. & Epstein, P.N. Metallothionein prevents diabetes-induced deficits in cardiomyocytes by inhibiting reactive oxygen species production. *Diabetes* **52**, 777–783 (2003).
18. Stegmaier, K. *et al.* Gene expression–based high-throughput screening (GE-HTS) and application to leukemia differentiation. *Nat. Genet.* **36**, 257–263 (2004).
19. Peck, D. *et al.* A method for high-throughput gene expression signature analysis. *Genome Biol.* **7**, R61 (2006).
20. Kim, Y.K. *et al.* Relationship of stereochemical and skeletal diversity of small molecules to cellular measurement space. *J. Am. Chem. Soc.* **126**, 14740–14745 (2004).
21. Franz, A.K., Dreyfuss, P.D. & Schreiber, S.L. Synthesis and cellular profiling of diverse organosilicon small molecules. *J. Am. Chem. Soc.* **129**, 1020–1021 (2007).
22. Larsson, N.G. & Clayton, D.A. Molecular genetic aspects of human mitochondrial disorders. *Annu. Rev. Genet.* **29**, 151–178 (1995).
23. Clayton, D.A. Transcription of the mammalian mitochondrial genome. *Annu. Rev. Biochem.* **53**, 573–594 (1984).
24. Antonetti, D.A., Reynet, C. & Kahn, C.R. Increased expression of mitochondrial-encoded genes in skeletal muscle of humans with diabetes mellitus. *J. Clin. Invest.* **95**, 1383–1388 (1995).
25. Huang, X. *et al.* Insulin-regulated mitochondrial gene expression is associated with glucose flux in human skeletal muscle. *Diabetes* **48**, 1508–1514 (1999).
26. Heddi, A., Stepien, G., Benke, P.J. & Wallace, D.C. Coordinate induction of energy gene expression in tissues of mitochondrial disease patients. *J. Biol. Chem.* **274**, 22968–22976 (1999).
27. Graham, D.J. *et al.* Incidence of hospitalized rhabdomyolysis in patients treated with lipid-lowering drugs. *J. Am. Med. Assoc.* **292**, 2585–2590 (2004).
28. Dirks, A.J. & Jones, K.M. Statin-induced apoptosis and skeletal myopathy. *Am. J. Physiol.* **291**, C1208–C1212 (2006).
29. van Vliet, A.K., Negre-Aminou, P., van Thiel, G.C., Bolhuis, P.A. & Cohen, L.H. Action of lovastatin, simvastatin, and pravastatin on sterol synthesis and their antiproliferative effect in cultured myoblasts from human striated muscle. *Biochem. Pharmacol.* **52**, 1387–1392 (1996).
30. Frendin, T.J. & Swainson, C.P. Acute renal failure secondary to non-traumatic rhabdomyolysis following amoxapine overdose. *N. Z. Med. J.* **98**, 690–691 (1985).
31. Blessing, W. & Walsh, J.C. Myotonia precipitated by propranolol therapy. *Lancet* **309**, 73–74 (1977).

32. Davidson, B.K. Myositis associated with griseofulvin therapy. *Am. Fam. Physician* **52**, 1277 (1995).
33. Delobel, P. & Pradinaud, R. Rhabdomyolysis associated with pentamidine isethionate therapy for American cutaneous leishmaniasis. *J. Antimicrob. Chemother.* **51**, 1319–1320 (2003).
34. Rowinsky, E.K. *et al.* Phase I and pharmacologic study of paclitaxel and cisplatin with granulocyte colony-stimulating factor: neuromuscular toxicity is dose-limiting. *J. Clin. Oncol.* **11**, 2010–2020 (1993).
35. Aronson, J.K. Ed. *Meyler's Side Effects of Drugs* 15th edn. (Elsevier Science, Burlington, Massachusetts, USA, 2006).
36. Bliss, C.I. The toxicity of poisons applied jointly. *Ann. Appl. Biol.* **26**, 585–615 (1939).
37. Kelley, D.E., He, J., Menshikova, E.V. & Ritov, V.B. Dysfunction of mitochondria in human skeletal muscle in type 2 diabetes. *Diabetes* **51**, 2944–2950 (2002).
38. Houstis, N., Rosen, E.D. & Lander, E.S. Reactive oxygen species have a causal role in multiple forms of insulin resistance. *Nature* **440**, 944–948 (2006).
39. Lustbader, J.W. *et al.* ABAD directly links Aβ to mitochondrial toxicity in Alzheimer's disease. *Science* **304**, 448–452 (2004).
40. Li, W.L., Zheng, H.C., Bukuru, J. & De Kimpe, N. Natural medicines used in the traditional Chinese medical system for therapy of diabetes mellitus. *J. Ethnopharmacol.* **92**, 1–21 (2004).
41. Bonawitz, N.D., Clayton, D.A. & Shadel, G.S. Initiation and beyond: multiple functions of the human mitochondrial transcription machinery. *Mol. Cell* **24**, 813–825 (2006).
42. Wu, Z. *et al.* Mechanisms controlling mitochondrial biogenesis and respiration through the thermogenic coactivator PGC-1. *Cell* **98**, 115–124 (1999).
43. Mootha, V.K. *et al.* Errα and Gabpa/b specify PGC-1α-dependent oxidative phosphorylation gene expression that is altered in diabetic muscle. *Proc. Natl. Acad. Sci. USA* **101**, 6570–6575 (2004).
44. Valle, I., Alvarez-Barrientos, A., Arza, E., Lamas, S. & Monsalve, M. PGC-1alpha regulates the mitochondrial antioxidant defense system in vascular endothelial cells. *Cardiovasc. Res.* **66**, 562–573 (2005).
45. Freyssenet, D., Irrcher, I., Connor, M.K., DiCarlo, M. & Hood, D.A. Calcium-regulated changes in mitochondrial phenotype in skeletal muscle cells. *Am. J. Phys.* **286**, 1053–1061 (2004).
46. Caprio, S. *et al.* Improvement of metabolic control in diabetic patients during mebendazole administration: preliminary studies. *Diabetologia* **27**, 52–55 (1984).
47. Karbowski, M. *et al.* Opposite effects of microtubule-stabilizing and microtubule-destabilizing drugs on biogenesis of mitochondria in mammalian cells. *J. Cell Sci.* **114**, 281–291 (2001).
48. Mitchell, P. Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. *Nature* **199**, 144–148 (1961).
49. Lee, P.D., Sladek, R., Greenwood, C.M. & Hudson, T.J. Control genes and variability: absence of ubiquitous reference transcripts in diverse mammalian expression studies. *Genome Res.* **12**, 292–297 (2002).
50. Hieronymus, H. *et al.* Gene expression signature-based chemical genomic prediction identifies a novel class of HSP90 pathway modulators. *Cancer Cell* **10**, 321–330 (2006).

Corrigendum: Large-scale chemical dissection of mitochondrial function

Bridget K Wagner, Toshimori Kitami, Tamara J Gilbert, David Peck, Arvind Ramanathan, Stuart L Schreiber, Todd R Golub & Vamsi K Mootha
Nat. Biotechnol. 26, 343–351 (2008); published online 24 February 2008; corrected after print 8 July 2008

In the version of this article initially published, on p.348, column 2, paragraph 2, line 7, the following sentence was incorrect: "Statins block the synthesis of cholesterol—a precursor to ubiquinone...." It should have read "Statins block the synthesis of mevalonate, a precursor not only of cholesterol but also ubiquinone, ..." The error has been corrected in the HTML and PDF versions of the article.

Corrigendum: Isolation and directed differentiation of neural crest stem cells derived from human embryonic stem cells

Gabsang Lee, Hyesoo Kim, Yechiel Elkabetz, George Al Shamy, Georgia Panagiotakos, Tiziano Barberi, Viviane Tabar & Lorenz Studer
Nat. Biotechnol. 25, 1468–1475 (2007); published online 25 November 2007; corrected after print 8 July 2008

In the version of this article initially published, a reference was missing from the first paragraph. A new sentence and the reference (no. 6) have been added: "A recent study characterized neural crest differentiation from cloned bovine blastocysts via a neural rosette intermediate[6]. Other...." Subsequent references have been renumbered. The corrections have been made in the HTML and PDF versions of the article.

Erratum: Looking forward, looking back

Anonymous
Nat. Biotechnol. 26, 475 (2008); published online May 2008; corrected after print 13 June 2008

In the version of this article initially published, in paragraph 4, the generic name and ligand given for Avastin are incorrect. The correct generic name is bevacizumab and its target is VEGF (vascular endothelial growth factor). The error has been corrected in the HTML and PDF versions of the article.

Erratum: Is personalized medicine finally arriving?

Malorye Allison
Nat. Biotechnol. 26, 509–517 (2008); published May, 2008; corrected after print 8 July 2008

In the version of this article initially published, Table 1 (pp. 510–511) contained two errors. In the entry for Agendia, the product Mammaprint was described as providing information on chemotherapy options for breast cancer patients. In fact Mammaprint is a prognostic test. In the entry for Genomic Health, the product Oncotype Dx was described as providing information on breast cancer recurrence. Oncotype DX also provides information on the response to chemotherapy. The error has been corrected in the HTML and PDF versions of the article.

EXHIBIT B

ARTICLES

nature genetics

PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes

Vamsi K Mootha[1,2,3,10], Cecilia M Lindgren[1,4,10], Karl-Fredrik Eriksson[4], Aravind Subramanian[1], Smita Sihag[1], Joseph Lehar[1], Pere Puigserver[5], Emma Carlsson[4], Martin Ridderstråle[4], Esa Laurila[4], Nicholas Houstis[1], Mark J Daly[1], Nick Patterson[1], Jill P Mesirov[1], Todd R Golub[1,5], Pablo Tamayo[1], Bruce Spiegelman[5], Eric S Lander[1,6], Joel N Hirschhorn[1,7,8], David Altshuler[1,2,7,9,11] & Leif C Groop[4,11]

DNA microarrays can be used to identify gene expression changes characteristic of human disease. This is challenging, however, when relevant differences are subtle at the level of individual genes. We introduce an analytical strategy, Gene Set Enrichment Analysis, designed to detect modest but coordinate changes in the expression of groups of functionally related genes. Using this approach, we identify a set of genes involved in oxidative phosphorylation whose expression is coordinately decreased in human diabetic muscle. Expression of these genes is high at sites of insulin-mediated glucose disposal, activated by PGC-1α and correlated with total-body aerobic capacity. Our results associate this gene set with clinically important variation in human metabolism and illustrate the value of pathway relationships in the analysis of genomic profiling experiments.

Type 2 diabetes mellitus (DM2) affects over 110 million people worldwide and is a principal contributor to atherosclerotic vascular disease, blindness, amputation and kidney failure[1]. Defects in insulin secretion are observed early in individuals with maturity-onset diabetes of the young, a monogenic form of type 2 diabetes[2]; insulin resistance at tissues including skeletal muscle is a cardinal feature of individuals with fully developed DM2. Many molecular pathways have been implicated in the disease process: β-cell development, insulin receptor signaling, carbohydrate production and utilization, mitochondrial metabolism, fatty acid oxidation, cytokine signaling, adipogenesis, adrenergic signaling and others. It is unclear, however, which of these or other pathways are disturbed in, and might be responsible for, DM2 in its common form.

Expression profiling using DNA microarrays enables researchers to survey the genome for transcripts whose levels are altered in tissue from individuals with disease. Microarray data can be used to classify individuals according to molecular characteristics and to generate hypotheses about disease mechanisms. This approach has been successful in the study of cancer[3], where large changes in the expression of individual genes have often been observed. When alterations in gene expression are more modest, however, the large number of genes tested, high variability between individuals and limited sample sizes typical of human studies make it difficult to distinguish true differences from noise.

One promising approach to increase power exploits the idea that alterations in gene expression might manifest at the level of biological pathways or coregulated gene sets, rather than individual genes. Subtle but coordinated changes in expression might be detected more readily by combining measurements across multiple members of each gene set. A straightforward strategy for identifying such differences is to examine top-ranking genes in a microarray experiment and then to create hypotheses about pathway membership. This is both subjective and *post hoc*, however, and thus prone to bias. A more objective set of approaches[4,5] tests for enrichment of pathway members among the top-ranking genes in a microarray study, comparing them to a null distribution in which genes are randomly distributed. Because functionally related genes are often coregulated, however, a positive result in such a test can be due solely to intrinsic correlation in gene expression rather than any relationship between expression of pathway members and the phenotype of interest.

We present an analytical technique designed to test *a priori* defined gene sets (for example, pathways) for association with disease phenotypes. We apply this method to gene expression profiles of human diabetic muscle, identifying a set of genes whose expression is correlated with insulin resistance and aerobic capacity. These results suggest hypotheses about pathways contributing to human metabolic disease and, more generally, show the value of incorporating information about functional relationships among genes in the analysis of microarray data.

[1]Whitehead Institute/MIT Center for Genome Research, Cambridge, Massachusetts, USA. [2]Department of Medicine, Harvard Medical School, Boston, Massachusetts, USA. [3]Department of Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA. [4]Department of Endocrinology, Wallenberg Laboratory, University Hospital MAS, Lund University, S-205 02 Malmo, Sweden. [5]Dana Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts, USA. [6]Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA. [7]Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA. [8]Divisions of Pediatrics and Endocrinology, Children's Hospital, Boston, Massachusetts, USA. [9]Department of Molecular Biology and Diabetes Unit, Massachusetts General Hospital, Boston, Massachusetts 02114, USA. [10]These authors contributed equally to this work. [11]These two authors contributed equally to this work. Correspondence should be addressed to D.A. (altshuler@molbio.mgh.harvard.edu) or L.C.G. (leif.groop@endo.mas.lu.se).

RESULTS

We used DNA microarrays to profile expression of over 22,000 genes in skeletal muscle biopsy samples from 43 age-matched males (**Table 1**), 17 with normal glucose tolerance (NGT), 8 with impaired glucose tolerance (IGT) and 18 with DM2. We obtained samples at the time of diagnosis (before treatment with hypoglycemic medication) and under the controlled conditions of a hyperinsulinemic euglycemic clamp. When assessed with either of two different analytical techniques[3,6] that take into account the multiple comparisons implicit in microarray analysis, no single gene had a significant difference in expression between the diagnostic categories (data not shown). This result is consistent with smaller studies[7,8] that did not identify any individual gene whose expression difference was significant when corrected for the large number of hypotheses tested[9,10].

Gene Set Enrichment Analysis

To test for sets of related genes that might be systematically altered in diabetic muscle, we devised a simple approach called Gene Set Enrichment Analysis (GSEA), which we introduce here (**Fig. 1**) and describe in more detail elsewhere (A.S. *et al.*, manuscript in preparation). The method combines information from the members of previously defined sets of genes (for example, biological pathways) to increase signal relative to noise and improve statistical power.

For a given pairwise comparison (for example, highly expressed in individuals with NGT versus those with DM2), we rank all genes according to the difference in expression (using an appropriate metric, such as signal-to-noise ratio, SNR). The null hypothesis of GSEA is that the rank ordering of the genes in a given comparison is random with regard to the diagnostic categorization of the samples. The alternative hypothesis is that the rank ordering of the pathway members is associated with the specific diagnostic criteria used to categorize the groups of affected individuals.

We then measure the extent of association by a non-parametric, running sum statistic termed the enrichment score (ES) and record the maximum ES (MES) over all gene sets in the actual data from affected individuals (**Fig. 1**). To assess the statistical significance of the MES, we use permutation testing of the diagnostic labels of the individuals (for example, whether an individual is affected with NGT or DM2; Fig. 1). Specifically, we compare the MES achieved in the actual data to that seen in each of 1,000 permutations that shuffled the diagnostic labels among the samples. The significance of the MES score is calculated as the fraction of the 1,000 random permutations in which the top pathway gave a stronger result than that observed in the actual data. Because the permutation test involves randomization of the diagnostic labels, it is a test for the dependence on the actual diagnostic status of the affected individuals. Moreover, because the actual MES is compared to the distribution of maximal ES values over all pathways examined in each of the randomized data sets, it accounts for multiple pathways tested, and no further correction is required[9,10].

Decreased expression of genes involved in oxidative phosphorylation

We applied GSEA to the microarray data described above, using 149 gene sets that we compiled (**Supplementary Table 1** online). Of these gene sets, 113 are grouped according to involvement in metabolic pathways (derived from public or local curation[11]) and 36 consist of gene clusters that are coregulated in a mouse expression atlas of 46 tissues[12]. The gene sets were selected without regard to the results of the microarray data from the affected individuals. The top gene set in

Table 1 Clinical and biochemical characteristics of male subjects with NGT, IGT and DM2

	Class			*P* value		
	NGT	IGT	DM2	NGT versus IGT	IGT versus DM2	NGT versus DM2
n	17	8	18			
Age, y	66.1 (1.0)	66.4 (1.6)	65.5 (1.8)			
BMI, kg/m^2	23.6 (3.4)	27.1 (4.8)	27.3 (4.0)			5.70×10^{-3}
WHR	0.91 (0.09)	0.97 (0.04)	0.99 (0.03)	3.00×10^{-2}		3.83×10^{-3}
Triglycerides, mmol/l	1.03 (0.40)	1.83 (1.60)	2.04 (1.13)			2.63×10^{-3}
Cholesterol, mmol/l	5.39 (0.09)	4.60 (1.48)	5.77 (0.97)			
OGTT						
Glucose 0 min, mmol/l	4.67 (0.50)	5.05 (0.46)	7.83 (2.3)		9.22×10^{-5}	2.01×10^{-5}
Insulin 0 min, μU/ml	5.41 (3.3)	13.38 (8.9)	12.0 (6.0)	4.05×10^{-2}		4.10×10^{-4}
Glucose 120 min, mmol/l	6.58 (0.94)	9.15 (0.8)	14.9 (4.0)	2.51×10^{-6}	8.91×10^{-6}	4.90×10^{-8}
Insulin 120 min, μU/ml	33.5 (19.3)	125.1 (66.1)	43.5 (25.6)	5.47×10^{-3}	9.73×10^{-3}	
M value, mg/kg/min	8.74 (3.15)	6.32 (3.08)	4.22 (1.72)			2.30×10^{-5}
VO2max, ml O_2/kg/min	32.1 (5.46)	26.5 (4.6)	24.3 (5.6)	1.72×10^{-2}		3.09×10^{-4}
Glycogen, mmol/kg	371.1 (77.0)	326.5 (88.0)	350.6 (97.8)			
Type I fibers						
Number, %	37.2 (13.5)	33.5 (3.6)	36.4 (9.3)			
Area, %	39.1 (14.4)	32.7 (0.91)	40.1 (10.7)		2.35×10^{-2}	
Capillaries/fiber	3.91 (0.72)	4.05 (1.04)	4.14 (0.75)			
Type IIb fibers						
Number, %	73.8 (42.1)	60.2 (51.4)	72.2 (36.7)			
Area, %	31.3 (18.0)	24.7 (18.3)	36.2 (15.4)			
Capillaries/fiber	2.97 (0.71)	3.05 (0.87)	3.02 (0.65)			

Values are mean (s.d.). OGTT, oral glucose tolerance test. M value is the total body glucose uptake. VO2max is the total body aerobic capacity. Only values of $P < 0.05$ are shown for pairwise comparisons, using a two-sided *t*-test.

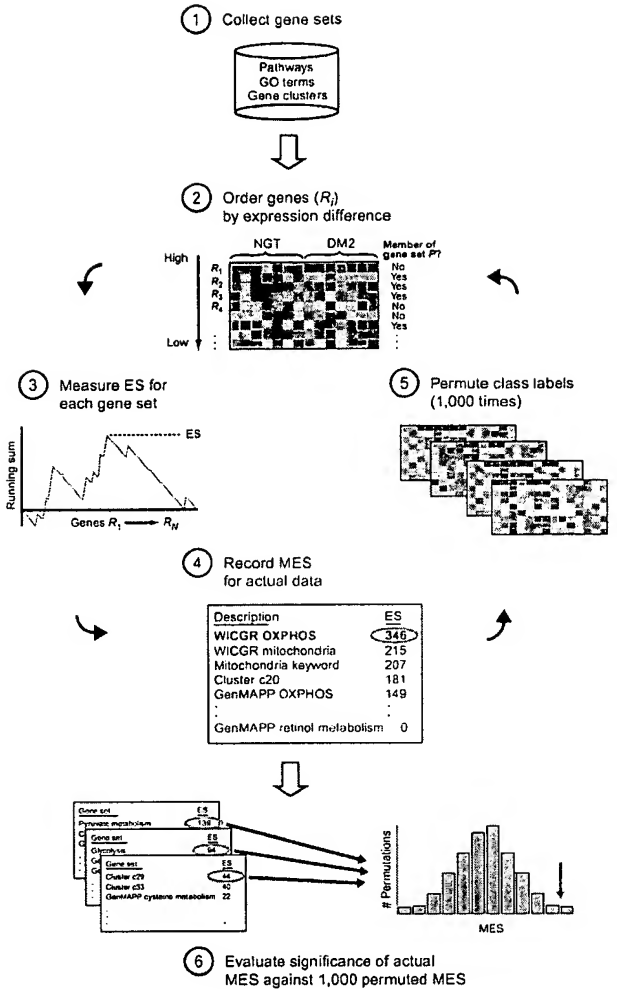


Figure 1 Schematic overview of GSEA. The goal of GSEA is to determine whether any *a priori* defined gene sets (step 1) are enriched at the top of a list of genes ordered on the basis of expression difference between two classes (for example, highly expressed in individuals with NGT versus those with DM2). Genes $R_1,...R_N$ are ordered on the basis of expression difference (step 2) using an appropriate difference measure (for example, SNR). To determine whether the members of a gene set *S* are enriched at the top of this list (step 3), a Kolmogorov-Smirnov (K-S) running sum statistic is computed: beginning with the top-ranking gene, the running sum increases when a gene annotated to be a member of gene set *S* is encountered and decreases otherwise. The ES for a single gene set is defined as the greatest positive deviation of the running sum across all *N* genes. When many members of *S* appear at the top of the list, ES is high. The ES is computed for every gene set using actual data, and the MES achieved is recorded (step 4). To determine whether one or more of the gene sets are enriched in one diagnostic class relative to the other (step 5), the entire procedure (steps 2–4) is repeated 1,000 times, using permuted diagnostic assignments and building a histogram of the maximum ES achieved by any pathway in a given permutation. The MES achieved using the actual data is then compared to this histogram (step 6, red arrow), providing us with a global *P* value for assessing whether any gene set is associated with the diagnostic categorization.

GSEA analysis yielded an MES score (MES = 346) that was significant at $P = 0.029$ over the 1,000 permutations of the 149 pathways. That is, in only 29 of 1,000 permutations did the top pathway (of the 149) exceed the score achieved by the top pathway achieved using the actual diagnostic labels.

The maximal ES score was obtained for an internally curated set consisting of genes involved in oxidative phosphorylation (we refer to this gene set as OXPHOS). Notably, the four gene sets with the next highest ES scores overlap with this OXPHOS gene set, and their enrichment is almost entirely explained by the overlap: a locally curated set of genes involved in mitochondrial function, a set of genes identified with the keyword 'mitochondria,' a cluster (referred to here as c20) of coregulated genes derived from the comparison of publicly available mouse data and a set of genes related to oxidative phosphorylation defined at the Affymetrix website[11].

Examination of the individual expression values for the 106 OXPHOS genes identifies the source of this signal (**Fig. 2**). Although the typical decrease in expression for individual OXPHOS genes is very modest (~20%), the decrease is consistent across the set: 89% (94 of 106) of the genes showing lower expression in individuals with DM2 relative to those with NGT (**Fig. 2**). As controls, we confirmed that the result is independent of specific aspects of data processing (such as scaling, thresholding, filtering) or of selection of difference metrics (data not shown). Moreover, the result identified by GSEA is supported by previous observations: others have suggested that oxidative capacities are altered in insulin resistant muscle[13,14], and recent microarray analyses of human diabetic muscle have identified genes in oxidative phosphorylation among their top-ranked genes (ref. 7 and M.E. Patti *et al.*, manuscript submitted).

OXPHOS-CR: a coregulated subset of OXPHOS genes

One of the overlapping gene sets identified by GSEA is cluster c20, defined as a set of genes that are tightly coregulated across many tissues. The partial overlap of OXPHOS with the coregulated cluster led us to ask whether all OXPHOS genes are coordinately regulated. We examined transcriptional coregulation of mouse homologs of OXPHOS genes across a mouse tissue expression atlas[12]. This identified a previously unrecognized subset of the OXPHOS biochemical pathway, corresponding to about two-thirds of the OXPHOS genes, that are strongly correlated across mouse tissues ($r = 0.61$; **Fig. 3a**). We term this subset OXPHOS-CR (oxidative phosphorylation co-regulated). The remaining OXPHOS genes show little co-regulation with OXHPOS-CR genes or with each other (**Fig. 3a**). The OXPHOS-CR subset was strongly expressed in 3 of 46 tissues: skeletal muscle, heart and brown fat. We note that these are the principal sites of insulin-mediated glucose disposal in mice.

We next asked whether the downregulation of OXPHOS observed in DM2 was a general property of all OXPHOS genes or was specific to OXPHOS-CR genes. Notably, the bulk of the statistical signal we observe in GSEA is accounted for by OXPHOS-CR (**Supplementary Fig. 1** online). Namely, the OXPHOS-CR subset showed a stronger mean deviation than the remainder of the OXPHOS gene set (mean SNR of 0.235 versus 0.128; $P = 0.04$) and was itself significant in the GSEA analysis (nominal *P* value = 0.001, as compared with nominal $P = 0.226$ for the remainder of the OXPHOS set). To see if these changes were secondary to hyperglycemia *per se* or preceded the onset of frank diabetes, we compared expression of OXPHOS-CR in individuals with NGT to that in individuals with IGT, the pre-diabetic state. We found that expression of OXPHOS-CR was also downregulated in individuals with IGT (nominal $P < 10^{-4}$). This suggests that downregulation of OXPHOS-CR precedes onset of hyperglycemia. Thus, GSEA allowed us to detect a subset of OXPHOS genes, called OXPHOS-CR, with three key properties: (i) they are members of the oxidative phosphorylation pathway (ii) they are tightly coregulated

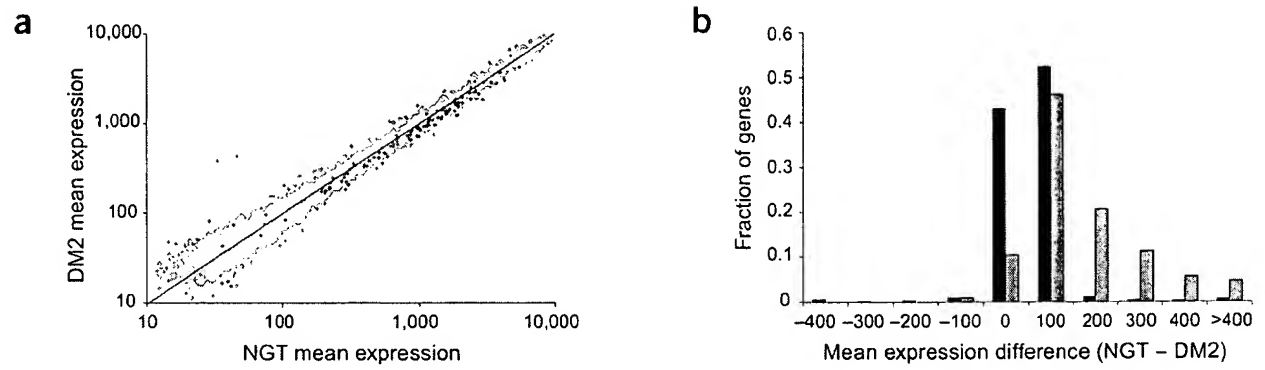


Figure 2 OXPHOS gene expression is reduced in diabetic muscle. (**a**) The mean expression of all genes (gray) and of OXPHOS genes (red) is plotted for individuals with DM2 versus those with NGT. (**b**) Histogram of mean gene expression level differences between individuals with NGT and DM2, using the data from **a**, for all genes (black) and for OXPHOS genes (red).

across many tissues and are highly expressed in the principal sites of insulin-mediated glucose disposal and (iii) their expression is subtly but consistently lower in muscle from individuals with both the pre-diabetic state IGT and DM2.

PGC-1α induces expression of OXPHOS-CR

The strong correlation in expression of the OXPHOS-CR genes and their coordinated downregulation in diabetic muscle led us to explore mechanisms that might mediate this tight control. We reasoned that peroxisome proliferator-activated receptor γ coactivator 1α (PGC-1α, encoded by *PPARGC1*), a cold-inducible regulator of mitochondrial biogenesis, thermogenesis and skeletal muscle fiber–type switching[15–17], was a prime candidate for mediating these effects. Consistent with this hypothesis, we observed that mean levels of *PPARGC1* transcript were similarly lower (by ~20%) in the diabetic muscle and noted that the promoters of several of the OXPHOS-CR genes have been reported to contain binding sites for nuclear respiratory factor 1, a transcription factor coactivated by PGC-1α[18].

To test directly whether OXPHOS-CR genes might be transcriptional targets of PGC-1α, we expressed PGC-1α in a mouse skeletal muscle cell line using an adenoviral expression vector[17] and used DNA microarrays to profile expression of the OXPHOS genes over a 3-d period. We found that a subset of OXPHOS genes was strongly upregulated in a time-dependent manner in response to PGC-1α and that this subset corresponded almost precisely to OXPHOS-CR (Fig. 3b). These *in vitro* results support the hypothesis that PGC-1α has a role in the regulation of OXPHOS-CR, both across the mouse tissue compendium and in the observed downregulation in diabetes.

Expression of OXPHOS-CR and measures of whole-body physiology

Metabolic control theory suggests that small adjustments in many sequential steps of a metabolic pathway can lead to a substantial change in the total flux through the pathway, whereas large changes in a single enzyme might have no measurable effects[19]. To test the hypothesis that differences in OXPHOS-CR gene expression in diabetic individuals might be related to changes in total body metabolism, we examined the relationships between diabetes status, expression of OXPHOS-CR genes and maximal oxygen uptake (VO2max) as measured in affected individuals (Fig. 4). Consistent

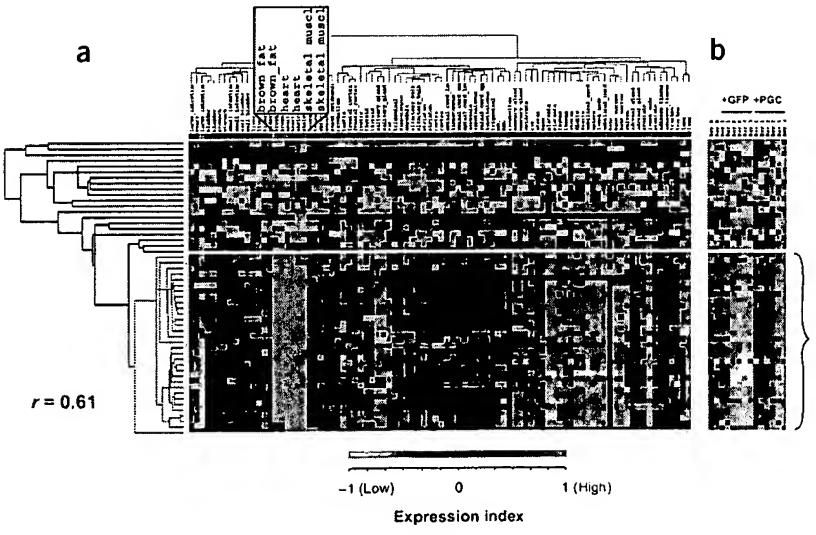


Figure 3 OXPHOS-CR represents a coregulated subset of OXPHOS genes responsive to the transcriptional coactivator PGC-1α. (**a**) Normalized expression profile of 52 mouse homologs of the human OXPHOS genes across the mouse expression atlas[12]. These 52 genes were hierarchically clustered[32]. The pink tree on the left corresponds to a subcluster with a correlation coefficient of 0.65. We call the human homologs of these mouse genes the OXPHOS-CR set. The human homologs of this tightly coregulated cluster, marked with an asterisk and delimited with a yellow box, are *ATP5J, ATP5L, ATP5O, COX5B, COX6A2, COX7A1, COX7B, COX7C, CYC1, CYCS, GRIM19, HSPC051, NDUFA2, NDUFA5, NDUFA7, NDUFA8, NDUFB3, NDUFB5, NDUFB6, NDUFC1, NDUFS2, NDUFS3, NDUFS5, SDHA, SDHB, UQCRB* and *UQCRC1*. (**b**) Normalized expression profile of OXPHOS mouse homologs in a mouse skeletal muscle cell line during a 3-d time course in response to PGC-1α. The expression profile includes infection with control vectors (expressing GFP) or with vectors expressing PGC-1α before infection (d 0) and 1, 2 and 3 d after adenoviral infection, all done in duplicate.

with previous reports[20], diabetes and VO2max were correlated in affected individuals (R^2_{adj} = 0.28, P = 0.0005). Notably, we found that the expression of OXPHOS-CR genes in muscle was strongly correlated with VO2max (R^2_{adj} = 0.22, P = 0.0012; **Fig. 4**), a measure of total-body physiology. The top ranking OXPHOS-CR gene, ubiquinol cytochrome c reductase binding protein (*UQCRB*), was even a stronger predictor (R^2_{adj} =0.31, P < 0.0001). Expression of OXPHOS-CR genes is not merely a proxy for diabetes status, however, because a two-variable regression of VO2max on diabetes status and OXPHOS-CR expression level shows that both variables contribute significantly to the correlation (P = 0.05 for the model with both variables as compared to the model with only diabetes status).

These results do not seem to be secondary to other known predictors of oxidative capacity. We found no relationship between body mass index or waist-to-hip ratio and OXPHOS-CR gene expression (R^2_{adj} < 0.01 in both cases). In addition, there was no significant relationship between quantitative measures of fiber types and OXPHOS-CR expression (data not shown). Thus, decreased in expression of OXPHOS-CR genes in muscle seems to be associated with changes in total-body aerobic capacity, even beyond their correlation to diabetes status, body habitus or muscle-fiber type.

DISCUSSION

Our results indicate that decreases in expression of OXPHOS-CR genes accompany, and might possibly contribute to, DM2. The relationship between OXPHOS and DM2 is richly supported by clinical investigation, exercise physiology, pharmacology and genetics. For example, the mitochondria of diabetic individuals show ultrastructural changes as well as decreases in oxidative phosphorylation activity[13,21]. Whole-body V02max (which we have shown to be correlated with OXPHOS-CR expression) predicts future development of DM2 (ref. 20). Exercise and caffeine consumption both increase oxidative phosphorylation capacity and can delay or prevent onset of diabetes[17,20,22,23]. Inherited mutations in mitochondrial DNA, which encodes 13 subunits of the electron transport chain, and whose copy number is under the control of PGC-1α[16], cause rare, inherited forms of diabetes[24]. Missense variants in PGC1-α have been reported to be associated with DM2 (refs. 25,26), although it is not yet clear if this association is reproducible[27]. Moreover, of the handful of genes in which variants have been clearly shown to influence risk of human diabetes, two are transcriptional partners of PGC1-α: HNF4-α (mutations of which cause early-onset diabetes) and PPARG, in which the Pro12Ala polymorphism is associated with risk of DM2 (reviewed in ref. 24). Further investigation will be required to test the hypothesis that the PGC-1α-regulated, OXPHOS-CR genes might represent a common link to these varied phenomena. If this hypothesis is valid, it would suggest that modulation of OXPHOS-CR activity might represent a target for the prevention and treatment of DM2.

More generally, methods like GSEA may be valuable in efforts to relate genomic variation to disease and measures of total-body physiology. Single-gene methods are powerful only when the individual gene effect is marked and the variance is small across individuals, which may not be the case in many disease states. Methods like GSEA are complementary to single-gene approaches and provide a framework with which to examine changes operating at a higher level of biological organization. This may be needed if common, complex disorders typically result from modest variation in the expression or activity of multiple members of a pathway. As gene sets are systematically assembled using functional and genomic approaches, methods such as GSEA will be valuable in detecting coordinated variation in gene function that contributes to common human diseases.

a

	Predictor(s)	R^2_{adj}	P value
	Diabetes status	0.28	0.0006
1	OXPHOS-CR	0.22	0.0012
1	Diabetes status, OXPHOS-CR	0.33	0.0004
2	*UQCRB*	0.31	<0.0001
2	Diabetes status, *UQCRB*	0.38	0.0001

[1] Addition of OXPHOS-CR improves the model with P = 0.05.

[2] Addition of *UQCRB* improves the model with P = 0.03.

b

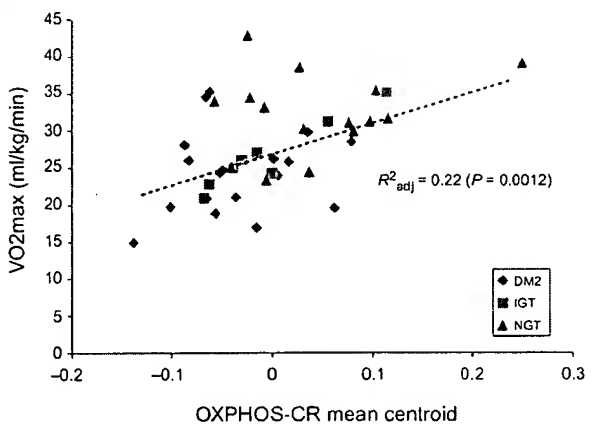


Figure 4 OXPHOS-CR predicts total-body aerobic capacity (VO2max). (a) Linear regression was used to model VO2max with diabetes status, the mean centroid of OXPHOS-CR gene expression, expression of *UQCRB* or in combination as explanatory (predictor) variables. The explanatory power and significance of the model are shown in the table. (b) Linear regression of VO2max against the mean centroid of OXPHOS-CR gene expression.

METHODS

Human subjects and clinical measurements. We selected 54 men of similar age but with varying degree of glucose tolerance who had been participating in The Malmö Prevention Study in southern Sweden for more than 12 years[20]. The investigation was approved by the Ethics Committee at Lund University, and informed consent was obtained from each of the volunteers. All subjects were Northern Europeans, and their glucose tolerance status was assessed using standardized 75-gram oral glucose tolerance test (OGTT) and by applying WHO85 criteria[20]. At the initial OGTT done 10 years earlier, none of the men had DM2 (ref. 20). An OGTT done at the time of the biopsy showed that 20 of the subjects had developed DM2, 8 fulfilled the criteria for IGT and 26 had NGT. As diabetes was diagnosed at the time of the repeat OGTT, none of the subjects were on medication for hyperglycemia or diabetes-related conditions.

Anthropometric and insulin sensitivity measures were done as previously described[28]. We measured height, weight, waist-to-hip ratio and fat-free mass on the day of the euglycemic clamp. We measured VO2max using an incremental work-conducted upright exercise test with a bicycle ergometer (Monark Varberg) combined with continuous analysis of expiratory gases and minute ventilation. Exercise was started at a workload varying between 30W and 100W depending on the previous history of endurance training or exercise habits and then increased by 20–50W every 3 min until a perceived exhaustion or a respiratory quotient of 1.0 was reached. Maximal aerobic capacity was defined as the VO2 during the last 30 s of exercise and is expressed per lean body mass. We determined insulin sensitivity with a standard 2-h euglycemic hyperinsulinemic clamp combined with infusion of tritiated glucose to estimate endogenous glucose production and indirect calorimetry (Deltatrac, Datex Instrumentarium) to estimate substrate oxidation[28]. We calculated the rate of glucose uptake (also referred to as the M value) from the infusion rate of glucose and the residual rate of endogenous glucose production measured by the tritiated glucose tracer during the clamp.

We took percutaneous muscle biopsy samples (20–50 mg) from the vastus lateralis muscle under local anesthesia (1% lidocaine) after the 2-h euglycemic hyperinsulinemic clamp using a Bergström needle[29]. We determined fiber-type composition and glycogen concentration as previously described[30]. We quantified and calculated the fibers using the COMFAS image analysis system (Scan Beam).

Cell culture and adenoviral infection. We cultured mouse myoblasts (C2C12 cells) and differentiated them into myotubes as previously described[16]. After 3 d of differentiation, we infected them with an adenovirus expressing either green fluorescent protein (GFP) or PGC-1α as previously described[17].

mRNA isolation, target preparation and hybridization. We prepared targets from human biopsy or mouse cell lines as previously described[3] and hybridized them to the Affymetrix HG-U133A or MG-U74Av2 chip, respectively. We selected only those scans with 10% Present calls and a *GAPD* 3′/*GAPD* 5′ expression ratio <1.33. We obtained gene expression data for 54 human samples, but only 43 met these selection criteria; the analysis in this paper is limited to these 43 individuals.

Data scaling and filtering. We subjected human microarray data to global scaling to correct for intensity-related biases. For each scan, we binned all genes according to their expression intensity and recorded the median intensity of each to serve as a calibration curve for that scan. We then scaled the expression to the calibration curve of the scan from one individual with NGT (individual mm12), which we visually inspected and deemed high-quality, using a linear interpolation between the calibration points. We then filtered the 22,283 genes on the HG-U133A chip to eliminate genes that had extremely low expression. A previous study suggested that an Affymetrix average difference level of 100 corresponds to an extremely low level ('not expressed'; ref. 12). Therefore, we only considered genes for which there was at least a single measure (average difference) greater than 100. Of the 22,283 genes on the HG-U133A chip, 10,983 genes met this filtering criterion.

Single gene microarray analysis. We carried out microarray analysis to identify individual genes that are significantly different between diagnostic classes using two software packages. First, we carried out marker analysis as previously described using GeneCluster[3]. Significance of individual genes was tested by permutation of class labels (5,000 iterations). We used both the *t*-test and SNR difference metrics in these analyses, both yielding comparable results. Second, we used the software package SAM, using a $\Delta = 0.5$, to search for gene expression values significantly different between classes[6].

Compilation of gene sets. We analyzed 149 gene sets consisting of manually curated pathways and clusters defined by public expression compendia (**Supplementary Table 1** online). First, we used two different sets of metabolic pathway annotations. We manually curated genes belonging to the following pathways: free fatty-acid metabolism, gluconeogenesis, glycolysis, glycogen metabolism, insulin signaling, ketogenesis, pyruvate metabolism, reactive oxygen species homeostasis, Krebs cycle, oxidative phosphorylation (OXPHOS) and mitochondria, using standard textbooks, literature reviews and LocusLink. We also downloaded NetAFFX[11] annotations (October 2002) corresponding to GenMAPP metabolic pathways. To identify sets of coregulated genes, we used self-organizing maps to group the GNF mouse expression atlas into 36 clusters[12,31]. Genes in these 36 groups were converted to Affymetrix HG-U133A probe sets using the ortholog tables available at the NetAFFX website (October 2002).

Rationale for grouped gene analysis. Consider a microarray data set with samples in two categories, *A* and *B*. For the sake of simplicity, let the size of *A* and *B* each be *n*. Consider a gene set *S* for which the expression levels differ between samples of *A* and *B*. Model the data set so that the entry D_{ij} for gene *i* and sample *j* is normally distributed with mean μ_{ij} and standard deviation σ, where

$$\mu_{ij} = \begin{cases} 0, & i \notin S \\ +\alpha, & i \in S, j \in A \\ -\alpha, & i \in S, j \in B \end{cases} .$$

Then the SNR for an individual gene in *S* is proportional to

$$\frac{\alpha\sqrt{n}}{\sigma} .$$

Suppose, on the other hand, that we know *S* and add the expression levels for all genes in *S*. Then the SNR is proportional to

$$\frac{\alpha\sqrt{nM}}{\sigma} ,$$

where *M* is the number of genes in *S*. This increases the mean of our statistic (which is standard normal for the null hypothesis of no gene set association) by a factor of $\sqrt{M}$. If the noise is in fact correlated for genes of *S*, this reduces the benefit, but we can still expect a large gain. In practice we will not be able to select a gene set containing fully concordant expression levels, but as long as an appreciable fraction of our gene set has this property, we can expect a benefit from the grouped gene approach.

Gene Set Enrichment Analysis (GSEA). GSEA determines if the members of a given gene set are enriched among the most differentially expressed genes between two classes. First, the genes are ordered on the basis of a difference metric. The results presented in the current manuscript use the SNR difference metric, which is simply the difference in means of the two classes divided by the sum of the standard deviations of the two diagnostic classes. In general, other difference metrics can also be used.

For each gene set, we then make an enrichment measure called the ES, which is a normalized Kolmogorov-Smirnov statistic. Consider the genes $R_1, \ldots, R_N$ that are ordered on the basis of the difference metric between the two classes and a gene set *S* containing *G* members. We define

$$X_i = -\sqrt{\frac{G}{N-G}}$$

if R_i is not a member of *S*, or

$$X_i = \sqrt{\frac{N-G}{G}}$$

if R_i is a member of *S*.

We then compute a running sum across all *N* genes. The ES is defined as

$$\max_{1 \le j \le N} \sum_{i=1}^{j} X_i$$

or the maximum observed positive deviation of the running sum. ES is measured for every gene set considered. To determine whether any of the given gene sets shows association with the class phenotype distinction, we permute the class labels 1,000 times, each time recording the maximum ES over all gene sets. In this regard, we are testing a single hypothesis. The null hypothesis is that no gene set is associated with the class distinction.

In this manuscript, after identifying OXPHOS-CR as a subset of co-regulated OXPHOS genes, we tested it (a single gene set) for association with clinical status using GSEA. Because OXPHOS-CR is not independent of the OXPHOS set interrogated in the initial analysis, this cannot be viewed as an independent hypothesis. For this reason, these *P* values are explicitly marked as nominal *P* values.

GSEA has been implemented as a software tool for use with microarray data and will be presented in fuller detail, including a discussion of different varieties of multiple hypothesis testing and applications to other biomedical problems, in a companion paper (A.S. *et al.*, manuscript in preparation).

Evaluating OXPHOS coregulation in mouse expression data sets. We used the NetAFFX website to identify probe sets on the mouse expression chips corresponding to human OXPHOS probe sets. We identified a total of 114 (106 of which passed our filtering criterion) probe sets corresponding to the human genes related to oxidative phosphorylation. Using the October 2002

ortholog tables at NetAFFX, we identified 61 mouse orthologs on the Affymetrix MG-U74Av2 chip. Of these 61 probe sets, 52 were represented in the GNF mouse expression atlas[12]. These expression data were normalized to a mean of 0 and a variance of 1. Data were hierarchically clustered and visualized using the Cluster and TreeView software packages[32]. We parsed these 52 genes into 32 coregulated probe sets and 20 probe sets that are not coregulated, based on the dendrogram in **Figure 3**. Forty distinct HG-U133A probe sets mapped to the 32 coregulated mouse probe sets, and 19 distinct HG-U133A probe sets mapped to the 20 mouse probe sets that are not coregulated. Five HG-U133A probe sets are shared between these two groups, representing ambiguous cases (human probe sets that map to two mouse probe sets, one of which is coregulated and the other of which is not). We omitted these five ambiguous human probe sets from our analysis. This left a total of 35 HG-U133A probe sets, which we call OXPHOS-CR genes, and a total of 14 HG-U133A probe sets, which we call OXPHOS not CR. Thirty-four and 13 of these genes, respectively, passed our filtering criteria, and these were used in **Supplementary Figure 1** online as well as in the OXPHOS-CR analysis described in the paper.

Linear regression analysis. We generated linear regression models using SAS (SAS Institute). We used clinical variables as dependent variables and OXPHOS-CR gene expression levels or other clinical/biochemical measures as the independent (explanatory or predictor) variables. To compute the mean centroid of OXPHOS-CR, we normalized the gene expression levels of the 34 OXPHOS-CR genes to a mean of 0 and a variance of 1 across all 43 individuals. The OXPHOS-CR mean centroid vector is simply the mean of these 34 expression vectors. In some regression analyses, we introduced dummy variables to represent diabetes status. For the regressions we carried out, we report the adjusted squared correlation coefficient (R^2_{adj}), which corrects for the degrees of freedom.

URLs. Further details on microarray data sets and analysis are available at http://www-genome.wi.mit.edu/mpg/oxphos/. Further data on microarrays are available at http://www-genome.wi.mit.edu/cancer/, http://www-stat.stanford.edu/~tibs/SAM/ and http://www.affymetrix.com/. The gene expression atlas is available at http://expression.gnf.org/.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank C. Ladd, M. Gaasenbeek and G. Ahlqvist for technical assistance; L. Gaffney for preparing illustrations; D. Stram and R. Heinrich for discussions; M. Patti and colleagues for sharing their manuscript before publication; B. Gewurz, E. Rosen, members of D.A. and E.S.L.'s labs for comments on the manuscript; and the individuals who volunteered for this study. V.K.M. is supported by a Howard Hughes Medical Institute physician postdoctoral fellowship. C.M.L. was supported by the Foundation for Strategic Research, the Royal Physiographic Society, the Sven Lundgrens Foundation and the Albert Pahlssons Foundation. T.R.G. is an Investigator of the Howard Hughes Medical Institute. J.N.H. is the recipient of a Career Development Award of the Burroughs Welcome Fund. D.A. is a Clinical Scholar in Translational Research of the Burroughs Welcome Fund and a Charles E. Culpeper Scholar of the Rockefeller Brothers Fund. This work was supported in part by grants from Affymetrix, Millennium Pharmaceuticals and Bristol-Myers Squibb to E.S.L. and from the Sigrid Juselius Foundation, the Juvenile Diabetes Foundation-Wallenberg Foundation, the Swedish Medical Research Council, the Novo-Nordisk Foundation and a European Community Genomics Integrated Force for Type 2 Diabetes grant to L.C.G.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 24 January; accepted 23 May 2003
Published online 15 June 2003; doi:10.1038/ng1180

1. Zimmet, P. Globalization, coca-colonization and the chronic disease epidemic: can the Doomsday scenario be averted? *J. Intern. Med.* **247**, 301–310 (2000).
2. Fajans, S.S., Bell, G.I. & Polonsky, K.S. Molecular mechanisms and clinical pathophysiology of maturity-onset diabetes of the young. *N. Engl. J. Med.* **345**, 971–980 (2001).
3. Golub, T.R. *et al.* Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* **286**, 531–537 (1999).
4. Doniger, S.W. *et al.* MAPPFinder: using Gene Ontology and GenMAPP to create a global gene-expression profile from microarray data. *Genome Biol.* **4**, R7 (2003).
5. Draghici, S., Khatri, P., Martins, R.P., Ostermeier, G.C. & Krawetz, S.A. Global functional profiling of gene expression. *Genomics* **81**, 98–104 (2003).
6. Tusher, V.G., Tibshirani, R. & Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. USA* **98**, 5116–5121 (2001).
7. Sreekumar, R., Halvatsiotis, P., Schimke, J.C. & Nair, K.S. Gene expression profile in skeletal muscle of type 2 diabetes and the effect of insulin treatment. *Diabetes* **51**, 1913–1920 (2002).
8. Yang, X., Pratley, R.E., Tokraks, S., Bogardus, C. & Permana, P.A. Microarray profiling of skeletal muscle tissues from equally obese, non-diabetic insulin-sensitive and insulin-resistant Pima Indians. *Diabetologia* **45**, 1584–1593 (2002).
9. Kropf, S. & Lauter, J. Multiple tests for different sets of variables using a data-driven ordering of hypotheses, with an application to gene expression data. *Biometrical J.* **44**, 789–800 (2002).
10. Storey, J.D. A direct approach to false discovery rates. *J. R. Statist. Soc. B* **64**, 479–498 (2002).
11. Liu, G. *et al.* NetAffx: Affymetrix probesets and annotations. *Nucleic Acids Res.* **31**, 82–86 (2003).
12. Su, A.I. *et al.* Large-scale analysis of the human and mouse transcriptomes. *Proc. Natl. Acad. Sci. USA* **99**, 4465–4470 (2002).
13. Bjorntorp, P., Schersten, T. & Fagerberg, S.E. Respiration and phosphorylation of mitochondria isolated from the skeletal muscle of diabetic and normal subjects. *Diabetologia* **3**, 346–352 (1967).
14. Simoneau, J.A., Colberg, S.R., Thaete, F.L. & Kelley, D.E. Skeletal muscle glycolytic and oxidative enzyme capacities are determinants of insulin sensitivity and muscle composition in obese women. *FASEB J.* **9**, 273–278 (1995).
15. Puigserver, P. *et al.* A cold-inducible coactivator of nuclear receptors linked to adaptive thermogenesis. *Cell* **92**, 829–839 (1998).
16. Wu, Z. *et al.* Mechanisms controlling mitochondrial biogenesis and respiration through the thermogenic coactivator PGC-1. *Cell* **98**, 115–124 (1999).
17. Lin, J. *et al.* Transcriptional co-activator PGC-1α drives the formation of slow- twitch muscle fibres. *Nature* **418**, 797–801 (2002).
18. Scarpulla, R.C. Nuclear activators and coactivators in mammalian mitochondrial biogenesis. *Biochim. Biophys. Acta* **1576**, 1–14 (2002).
19. Brown, G.C. Control of respiration and ATP synthesis in mammalian mitochondria and cells. *Biochem. J.* **284**, 1–13 (1992).
20. Eriksson, K.F. & Lindgarde, F. Impaired glucose tolerance in a middle-aged male urban population: a new approach for identifying high-risk cases. *Diabetologia* **33**, 526–531 (1990).
21. Kelley, D.E., He, J., Menshikova, E.V. & Ritov, V.B. Dysfunction of mitochondria in human skeletal muscle in type 2 diabetes. *Diabetes* **51**, 2944–2950 (2002).
22. van Dam, R.M. & Feskens, E.J. Coffee consumption and risk of type 2 diabetes mellitus. *Lancet* **360**, 1477–1478 (2002).
23. Ojuka, E.O., Jones, T.E., Han, D.H., Chen, M. & Holloszy, J.O. Raising Ca^{2+} in L6 myotubes mimics effects of exercise on mitochondrial biogenesis in muscle. *FASEB J.* **17**, 675–681 (2003).
24. Florez, J.C., Hirschhorn, J.N. & Altshuler, D. The inherited basis of diabetes mellitus: general lessons for the genetic analysis of complex traits. *Annu. Rev. Genomics Hum. Genet.* (in the press).
25. Ek, J. *et al.* Mutation analysis of peroxisome proliferator-activated receptor-gamma coactivator-1 (PGC-1) and relationships of identified amino acid polymorphisms to Type II diabetes mellitus. *Diabetologia* **44**, 2220–2226 (2001).
26. Hara, K. *et al.* A genetic variation in the *PGC-1* gene and insulin resistance. *Diabetologia* **45**, 740–743 (2002).
27. Lacquemant, C., Chikri, M., Boutin, P., Samson, C. & Froguel, P. No association between the G482S polymorphism of the proliferator-activated receptor-γ coactivator-1 (*PGC-1*) gene and Type II diabetes in French Caucasias. *Diabetologia* **45**, 602–603; author reply 604 (2002).
28. Groop, L. *et al.* Metabolic consequences of a family history of NIDDM (the Botnia study): evidence for sex-specific parental effects. *Diabetes* **45**, 1585–1593 (1996).
29. Eriksson, K.F., Saltin, B. & Lindgarde, F. Increased skeletal muscle capillary density precedes diabetes development in men with impaired glucose tolerance. A 15-year follow-up. *Diabetes* **43**, 805–808 (1994).
30. Schalin-Jantti, C., Laurila, E., Lofman, M. & Groop, L.C. Determinants of insulin-stimulated skeletal muscle glycogen metabolism in man. *Eur. J. Clin. Invest.* **25**, 693–698 (1995).
31. Tamayo, P. *et al.* Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. *Proc. Natl. Acad. Sci. USA* **96**, 2907–2912 (1999).
32. Eisen, M.B., Spellman, P.T., Brown, P.O. & Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA* **95**, 14863–14868 (1998).

OPEN ACCESS Freely available online

PLoS one

Pioglitazone Enhances Mitochondrial Biogenesis and Ribosomal Protein Biosynthesis in Skeletal Muscle in Polycystic Ovary Syndrome

Vibe Skov[1]*, Dorte Glintborg[2], Steen Knudsen[3], Qihua Tan[1,4], Thomas Jensen[3], Torben A. Kruse[1], Henning Beck-Nielsen[2], Kurt Højlund[2]*

1 Department of Biochemistry, Genetics, and Pharmacology, Odense University Hospital and Human Microarray Centre (HUMAC), University of Southern Denmark, Odense, Denmark, **2** Diabetes Research Centre, Department of Endocrinology, Odense University Hospital, Odense, Denmark, **3** Medical Prognosis Institute Aps, Hørsholm, Denmark, **4** Institute of Public Health, University of Southern Denmark, Odense, Denmark

Abstract

Insulin resistance is a common metabolic abnormality in women with PCOS and leads to an elevated risk of type 2 diabetes. Studies have shown that thiazolidinediones (TZDs) improve metabolic disturbances in PCOS patients. We hypothesized that the effect of TZDs in PCOS is, in part, mediated by changes in the transcriptional profile of muscle favoring insulin sensitivity. Using Affymetrix microarrays, we examined the effect of pioglitazone (30 mg/day for 16 weeks) on gene expression in skeletal muscle of 10 obese women with PCOS metabolically characterized by a euglycemic-hyperinsulinemic clamp. Moreover, we explored gene expression changes between these PCOS patients before treatment and 13 healthy women. Treatment with pioglitazone improved insulin-stimulated glucose metabolism and plasma adiponectin, and reduced fasting serum insulin (all $P<0.05$). Global pathway analysis using Gene Map Annotator and Pathway Profiler (GenMAPP 2.1) and Gene Set Enrichment Analysis (GSEA 2.0.1), revealed a significant upregulation of genes representing mitochondrial oxidative phosphorylation (OXPHOS), ribosomal proteins, mRNA processing reactome, translation factors, and proteasome degradation in PCOS after pioglitazone therapy. Quantitative real-time PCR suggested that upregulation of OXPHOS genes was mediated by an increase in PGC-1α expression ($P<0.05$). Pretreatment expression of genes representing OXPHOS and ribosomal proteins was down-regulated in PCOS patients compared to healthy women. These data indicate that pioglitazone therapy restores insulin sensitivity, in part, by a coordinated upregulation of genes involved in mitochondrial OXPHOS and ribosomal protein biosynthesis in muscle in PCOS. These transcriptional effects of pioglitazone may contribute to prevent the onset of type 2 diabetes in these women.

Citation: Skov V, Glintborg D, Knudsen S, Tan Q, Jensen T, et al. (2008) Pioglitazone Enhances Mitochondrial Biogenesis and Ribosomal Protein Biosynthesis in Skeletal Muscle in Polycystic Ovary Syndrome. PLoS ONE 3(6): e2466. doi:10.1371/journal.pone.0002466

Editor: Cathal Seoighe, University of Cape Town, South Africa

Received April 2, 2008; **Accepted** May 14, 2008; **Published** June 18, 2008

Funding: This work was supported in part by grants from the Novo Nordisk Foundation, the Danish Diabetes Association, and the Danish Medical Research Council.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: vibe.skov@ouh.regionsyddanmark.dk (VS); k.hojlund@dadlnet.dk (KH)

Introduction

Polycystic ovary syndrome (PCOS) is a common endocrine and metabolic disorder occurring in 5–10% of premenopausal women. Symptoms of PCOS include menstrual irregularities, hyperandrogenism, and infertility [1]. Several studies have shown that insulin resistance plays an important role in the pathogenesis of PCOS [2] and increases the risk for development of type 2 diabetes [1,2]. Thiazolidinediones (TZDs), including pioglitazone, are peroxisome proliferator-activated receptor (PPAR)-γ agonists that induce adipogenesis and have insulin-sensitizing and antidiabetic properties [3]. TZDs operate via receptor-dependent or independent mechanisms thereby regulating the expression of genes involved in mitochondrial biogenesis, insulin signal transduction, and glucose and lipid metabolism [4,5]. PPARγ is most abundantly expressed in adipose tissue and to a lesser extent in muscle and liver tissue [6].

Recent studies have demonstrated that the beneficial metabolic effects of treatment with pioglitazone in PCOS patients partly occur through improvement of insulin sensitivity, including increased insulin-stimulated total, oxidative and non-oxidative glucose transport, and a decrease in insulin secretion [7,8], similar to observations in patients with type 2 diabetes [9]. Skeletal muscle accounts for the majority of insulin-stimulated glucose transport suggesting an important role for TZDs in this tissue.

The mechanisms by which TZDs exert their insulin sensitizing action in skeletal muscle of type 2 diabetic patients are not yet fully understood, but may include increased downstream insulin receptor signaling [10,11] and enhanced fatty acid uptake and oxidation [12,13]. Moreover, animal studies have demonstrated that prolonged treatment with TZDs is associated with increased AMP-activated protein kinase (AMPK) activity [14], and increased expression of NADH dehydrogenase subunit 1 of complex I and PPAR-γ coactivator-1α (PGC-1α) in skeletal muscle [15]. These observations suggest that TZDs may increase insulin sensitivity, in part, by improving mitochondrial oxidative metabolism (OXPHOS). We have recently demonstrated reduced expression of genes involved in OXPHOS in skeletal muscle of insulin resistant women with PCOS [16]. However, to our knowledge, no study

has examined whether changes in muscle transcripts contribute to the insulin sensitizing effect of TZDs in PCOS patients.

DNA microarrays are high throughput technologies aiming at measuring transcript abundance for thousands of genes simultaneously [17]. To explore mRNA levels in diseases where differences in gene expression are too modest or numerous to extract meaningful biological function from individual genes, the application of global pathway analysis has been promising in detecting coordinated changes in gene expression levels [16,18–20].

In the present study, we hypothesized that pioglitazone treatment may modify the expression of pathways representing OXPHOS genes as well as other pathways implicated in skeletal muscle insulin resistance of PCOS patients. We used two different approaches for global pathway analysis, and quantitative real-time PCR (q-RT-PCR) was applied to evaluate microarray results.

Results

Clinical and metabolic characteristics

After treatment with pioglitazone, basal serum insulin levels were reduced, whereas plasma adiponectin increased 2-fold (all $P<0.01$) (Table 1). No significant changes in plasma triglycerides, basal plasma glucose, serum free testosterone, or plasma FFA levels were observed. Insulin-stimulated total glucose disposal was increased by 36% in PCOS patients in response to pioglitazone ($P<0.01$), and this was partly accounted for by a 26% increase in glucose oxidation ($P<0.05$) and a 50% increase in non-oxidative glucose metabolism ($P<0.01$). There was no significant change in the ability of insulin to suppress lipid oxidation in response to pioglitazone.

Pretreatment levels of plasma triglycerides ($P<0.05$), basal serum insulin and free testosterone levels were elevated ($P<0.01$), and plasma adiponectin tended to be decreased ($P<0.07$) in PCOS patients compared with control subjects. There was no significant difference with respect to basal plasma glucose and FFA. Insulin-mediated total, non-oxidative and oxidative glucose metabolism and suppression of lipid oxidation were impaired in PCOS patients compared with control subjects (all $P<0.01$).

Oligonucleotide microarray analysis

Statistical analysis of the 54.675 probe sets represented on the array revealed that 23.933 probe sets were differentially regulated in PCOS patients in response to treatment with pioglitazone (uncorrected $P<0.05$). 5.303 probe sets remained significantly differentially expressed when corrected for multiple testing using the Benjamini-Hochberg method (FDR<0.01) [21].

Pretreatment levels of 14.834 probe sets were differentially expressed in PCOS patients compared with control subjects. After correction for multiple testing (FDR<0.05), 2.754 probe sets remained significantly differentially expressed [21]. To examine this huge amount of regulated transcripts, we decided to apply global pathway analysis to identify significantly regulated pathways.

Effects of pioglitazone evaluated by global pathway analysis

The pathways that were regulated in skeletal muscle of PCOS patients in response to pioglitazone treatment were examined by Gene Map Annotator and Pathway Profiler (GenMAPP 2.1) together with its accessory program, MAPPFinder 2.1, and by gene set enrichment analysis (GSEA 2.0.1). Using MAPPFinder, the top ranked upregulated pathways were *ribosomal proteins*, *electron transport chain*, *mRNA processing reactome*, *ubiquinone biosynthesis*, *proteasome degradation*, and *translation factors* (family wise error rate (FWER) <0.05) (Table 2). Applying GSEA, the *electron transport chain*, *VOXPHOS*, *insulin 2F up* (genes 2 fold upregulated by insulin), *ribosomal proteins*, *mRNA splicing*, *mRNA processing*, *circadian exercise* (genes that regulate the 24-hour cycle), *rapamycin_DN*, *glutamine_DN* (genes downregulated in response to rapamycin or glutamine

Table 1. Clinical and metabolic characteristics of PCOS patients and control subjects.

	Control Subjects	PCOS Pretreatment	PCOS Posttreatment
n	13	10	10
Age (years)	34.7±2.0	30.3±2.1	
Body mass index (kg/m^2)	34.0±1.8	33.2±0.9	33.0±1.1
Body fat (%)	41±1.6	39.1±1.3	39.8±1.4
Plasma triglycerides (mmol/l)	0.86±0.12 †	1.43±0.22	1.15±0.16
Serum free testosterone (nmol/l)	0.025±0.003 ††	0.053±0.009	0.048±0.007
Plasma glucose basal (mmol/l)	5.6±0.1	5.9±0.2	5.6±0.1
Serum insulin basal (pmol/l)	51±6 ††	125±23	69±12††
Plasma FFA basal (mmol/l)	0.49±0.04	0.45±0.05	0.41±0.05
Plasma adiponectin (mg/l)	9.4±0.8	7.1±0.9	16.1±2.2††
Rd basal (mg·min^{-1}·m^{-2})	72±3	75±3	76±4
Rd clamp (mg·min^{-1}·m^{-2})	290±23 ††	138±18	188±25††
Glucose oxidation basal (mg·min^{-1}·m^{-2})	52±8	46±5	43±10
Glucose oxidation clamp (mg·min^{-1}·m^{-2})	124±5 ††	80±10	101±12†
Lipid oxidation basal (mg·min^{-1}·m^{-2})	33±3	38±1	40±4
Lipid oxidation clamp (mg·min^{-1}·m^{-2})	7±2 ††	24±4	16±5
NOGD basal (mg·min^{-1}·m^{-2})	20±7	30±5	34±9
NOGD clamp (mg·min^{-1}·m^{-2})	165±22 ††	58±12	87±14††

PCOS pretreatment vs. controls, and the effect of pioglitazone treatment (30 mg/day for16 weeks) in PCOS patients. Students *T*-test for non-paired and paired data used, respectively. Data represent means±SEM. ††P<0.01 and †P<0.05 vs. PCOS pretreatment. NOGD, non-oxidative glucose disposal.

doi:10.1371/journal.pone.0002466.t001

Table 2. The top-ten most upregulated pathways analyzed with MAPPFinder 2.1.

MAPP Name	Changed (n)	Measured (n)	ON MAPP (n)	Changed (%)	Z Score	Permute p-value	FWER p-value
Ribosomal proteins	69	88	88	78.4	10.9	<0.0005	<0.0005
Electron transport chain	61	91	105	67.0	8.6	<0.0005	<0.0005
mRNA processing reactome	62	125	127	49.6	5.7	<0.0005	<0.0005
Ubiquinone biosynthesis	32	54	81	59.3	5.3	<0.0005	<0.0005
Proteasome degradation	31	60	61	51.7	4.3	<0.0005	0.003
Translation factors	26	50	50	52.0	4.0	<0.0005	0.02
TGF beta receptor netpath 7	57	151	151	37.7	3.0	0.003	0.43
Striated muscle contraction	18	38	38	47.4	2.8	0.007	0.56
Androgen receptor netpath 2	43	112	112	38.4	2.7	0.008	0.65
TNF alpha NFkB netpath 9	66	186	187	35.5	2.6	0.005	0.75

A p-value<0.05 and a fold change ≥1.05 were used as the criteria for gene expression changes in PCOS patients after pioglitazone treatment. The z-score is based on N = 4998 genes linked to a MAPP and R = 1355 of these genes meeting the criteria for change in expression. Changed (n): number of genes changed. Measured (n): number of genes measured on the chip. On MAPP (n): number of genes on the MAPP. Changed (%): Changed (n) divided by Measured (n). FWER p-value: Family Wise Error Rate.

doi:10.1371/journal.pone.0002466.t002

starvation, respectively), *mRNA processing reactome*, *oxidative phosphorylation*, *translation factors*, *Human MitoDB 6*, *proteasome degradation*, and *leucine_DN* (genes downregulated in response to leucine starvation) were significantly upregulated (FWER<0.05) (Table 3). A further description of gene sets from GSEA can be obtained from the GSEA homepage [22]. No pathways or gene sets were significantly downregulated (FWER<0.05) when using MAPPFinder or GSEA (Table S1 and S2).

Using GO, *ribonucleoprotein complex* and *ribosome* were the most upregulated cellular components (C), and *structural constituent of ribosome* was the most upregulated molecular function (F) term. The most upregulated biological process (P) term was *protein biosynthesis*. *Oxidoreductase activity, acting on NADH or NADPH and NADH dehydrogenase activity* was among the top twenty upregulated GO terms (Table S3). *Signal transducer activity* was the most downregulated molecular function (F) term, and *cell communication* was the most downregulated biological process (P) term. Membrane and related terms were the most downregulated cellular components (C) (Table S4).

Evaluating the results from both GenMAPP and GSEA, pathways representing OXPHOS genes, *ribosomal proteins*, *mRNA processing reactome*, *translation factors*, and *proteasome degradation* were significantly upregulated (FWER<0.05).

Pretreatment abnormalities in PCOS examined by global pathway analysis

To test whether the transcriptional changes induced by pioglitazone were correcting preexisting abnormalities in patients

Table 3. Ranking of the 15 most upregulated gene sets analyzed with GSEA 2.0.1.

Name	Size	ES	NES	NOM p-value	FDR q-value	FWER p-value
Electron transport chain	97	−0.60	−3.00	<0.0001	<0.0001	<0.0001
VOXPHOS	77	−0.62	−3.00	<0.0001	<0.0001	<0.0001
Insulin 2F up	188	−0.52	−2.89	<0.0001	<0.0001	<0.0001
Ribosomal proteins	88	−0.53	−2.61	<0.0001	<0.0001	<0.0001
mRNA splicing	48	−0.56	−2.43	<0.0001	<0.0001	<0.0001
mRNA processing	42	−0.57	−2.36	<0.0001	<0.0001	<0.0001
Circadian exercise	42	−0.56	−2.32	<0.0001	0.0002	0.0005
Rapamycin DN	189	−0.40	−2.19	<0.0001	0.0007	0.002
Glutamine DN	251	−0.38	−2.12	<0.0001	0.001	0.004
mRNA processing reactome	108	−0.40	−2.01	<0.0001	0.003	0.01
Oxidative phosphorylation	58	−0.44	−1.97	<0.0001	0.004	0.01
Translation factors	47	−0.46	−1.96	<0.0001	0.004	0.02
Human MitoDB 6 2002	385	−0.32	−1.95	<0.0001	0.005	0.02
Proteasome degradation	31	−0.50	−1.93	<0.0001	0.005	0.02
Leucin DN	141	−0.36	−1.92	<0.0001	0.005	0.03

Gene sets with a FWER<0.05 are shown. All genes on the chip were ranked by difference in expression after pioglitazone treatment of PCOS patients using the t-test. An enrichment score (ES) was assigned to each gene, and the maximum ES (MES) was calculated for each gene set. NES: Enrichment score normalized for differences in gene set size. FDR q-value: False Discovery Rate. FWER p-value: Family Wise Error Rate.

doi:10.1371/journal.pone.0002466.t003

Table 4. The ten most downregulated pathways analyzed with MAPPFinder 2.1.

MAPP Name	Changed (n)	Measured (n)	ON MAPP (n)	Changed (%)	Z Score	Permute p-value	FWER p-value
Ribosomal proteins	43	88	88	48.9	9.2	<0.0005	<0.0005
Electron transport chain	39	91	105	42.9	7.7	<0.0005	<0.0005
Aminoacyl tRNA biosynthesis	10	23	24	43.5	3.9	<0.0005	0.07
Ubiquinone biosynthesis	17	54	81	31.5	3.5	0.002	0.18
Phenylalanine, tyrosine, and tryptophan biosynthesis	5	11	36	45.5	2.9	0.01	0.55
Inositol metabolism	1	1	9	100.0	2.4	0.13	0.96
Translation factors	13	50	50	26.0	2.3	0.02	0.97
mRNA processing reactome	27	125	127	21.6	2.2	0.03	0.98
Cell cycle G1 to S control reactome	16	67	67	23.9	2.2	0.04	0.99
Mitochondrial fatty acid betaoxidation	5	16	16	31.3	1.9	0.08	1
Insulin signaling	31	159	159	19.5	1.8	0.07	1

A p-value<0.05 and a fold change ≤ -1.05 were used as the criteria for gene expression changes between PCOS patients and control subjects. The z-score is based on N = 4998 genes linked to a MAPP and R = 730 of these genes meeting the criteria for change in expression. Changed (n): number of genes changed. Measured (n): number of genes measured on the chip. On MAPP (n): number of genes on the MAPP. Changed (%): Changed (n) divided by Measured (n). FWER p-value: Family Wise Error Rate.

doi:10.1371/journal.pone.0002466.t004

with PCOS, MAPPFinder 2.1 and GSEA 2.0.1 were also applied for analysis of gene expression changes between 10 PCOS patients before treatment and 13 control subjects. Pretreatment muscle biopsies from seven of the 10 PCOS patients analyzed in the present study were also included in a previous microarray study, in which the 16 most insulin resistant of all larger cohort of PCOS patients [7] were compared with the same 13 control subjects [16]. In MAPPFinder, *ribosomal proteins* and the *electron transport chain* were significantly downregulated (FWER<0.05) (Table 4). Using GSEA, the expression of *VOXPHOS, electron transport chain, valine, leucine, and isoleucine degradation, propanoate metabolism, oxidative phosphorylation, Human mitoDB 6, insulin 2F up, ribosomal proteins,* and *fatty acid metabolism* was significantly decreased (FWER<0.05) (Table 5). No pathways and gene sets remained significantly upregulated in MAPPFinder and GSEA after correction for multiple testing (FWER) (Table S5 and S6). Applying GO, *cell communication* was the most upregulated biological process (P), *enzyme regulator activity* the most upregulated molecular function (F), and *intrinsic to membrane* the most upregulated cellular component (Table S7). The most downregulated GO terms were *ribosome* and *ribonucleoprotein complex* (C), *structural constituent of ribosome* (F), and *protein biosynthesis* (P) (Table S8).

Significant downregulation (FWER<0.05) of genes in pathways representing OXPHOS and ribosomal proteins in PCOS patients was identified using both GenMAPP and GSEA, and, in part, validated our previous findings, where we used the same cohort of control subjects and 16 insulin resistant PCOS patients [16].

Validation of pathway analysis with q-RT-PCR

The effect of pioglitazone treatment on expression of genes involved in OXPHOS and ribosomal protein synthesis and degradation was validated by q-RT-PCR (Table S9). All five respiratory genes examined and UCP2 showed an increase in expression (1.2–2.1 fold). Of these, NDUFA3 and SDHD were significantly upregulated, and ATP5H tended ($P = 0.057$) to be upregulated (Figure 1*A*). The expression of genes known to

Table 5. Ranking of the ten most downregulated gene sets analyzed with GSEA 2.0.1.

Name	Size	ES	NES	NOM p-value	FDR q-value	FWER p-value
VOXPHOS	77	0.61	3.06	<0.0001	<0.0001	<0.0001
Electron transport chain	97	0.58	2.99	<0.0001	<0.0001	<0.0001
Valine, leucine, and isoleucine degradation	36	0.55	2.30	<0.0001	<0.0001	<0.0001
Propanoate metabolism	31	0.57	2.29	<0.0001	<0.0001	<0.0001
Oxidative phosphorylation	58	0.46	2.13	<0.0001	0.0007	0.002
Human mitoDB 6 2002	385	0.34	2.13	<0.0001	0.0005	0.002
Insulin 2F up	188	0.35	2.05	<0.0001	0.002	0.005
Ribosomal proteins	88	0.42	2.04	<0.0001	0.002	0.007
Fatty acid metabolism	79	0.40	1.94	<0.0001	0.006	0.03
Cell cycle arrest	32	0.45	1.83	0.005	0.01	0.06

All genes on the chip were ranked by difference in expression between PCOS patients and control subjects using the t-test. An enrichment score (ES) was assigned to each gene, and the maximum ES (MES) was calculated for each gene set. NES: Enrichment score normalized for differences in gene set size. FDR q-value: False Discovery Rate. FWER p-value: Family Wise Error Rate.

doi:10.1371/journal.pone.0002466.t005

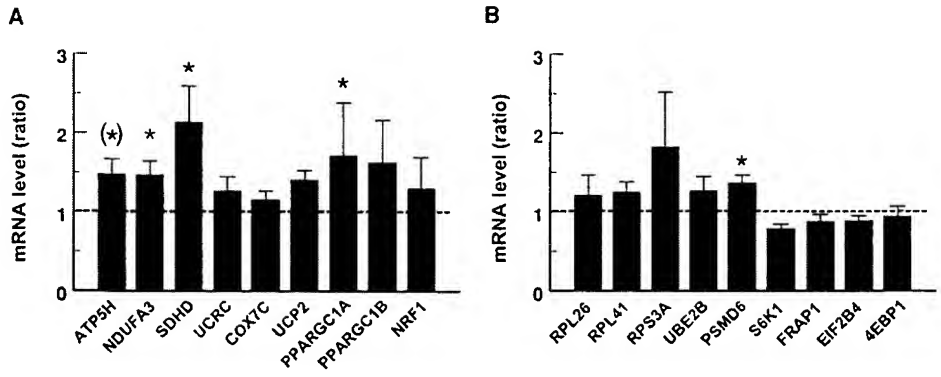


Figure 1. Effect of pioglitazone on expression of muscle genes. The mRNA expression level of selected genes was determined by quantitative real-time PCR in nine patients with the polycystic ovary syndrome (PCOS). The relative expression of genes involved in oxidative phosphorylation (*A*) and ribosomal protein biosynthesis including signaling to protein synthesis (*B*) are given. Regulated genes have mRNA levels different from 1.0 (dotted line). Data are means±SEM. *$P<0.05$.
doi:10.1371/journal.pone.0002466.g001

mediate transcriptional control of mitochondrial biogenesis, PPARGC1A (PGC-1α), PPARGC1B (PGC-1β) and NRF1, increased 1.4–1.7 fold in response to pioglitazone. However, only the increase in PGC-1α achieved statistical significance ($P<0.02$). The mRNA level of PSMD6 was significantly increased, whereas the increase (1.2–1.8 fold) in expression of the other genes representing protein metabolism, RPS3A, RPL26, RPL41, and UBE2B, was not significant (Figure 1*B*). Moreover, the gene expression of enzymes involved in signaling to protein synthesis including EIF2B4, FRAP1, S6K1 and 4EBP1 was unaltered in response to pioglitazone. We did not evaluate gene expression changes in skeletal muscle between patients with PCOS and control subjects with q-RT-PCR, because these changes have been validated recently [16].

Discussion

In this study, we examined the molecular signature of pioglitazone therapy in skeletal muscle of women with PCOS using global transcriptional profiling. We demonstrate that increased insulin sensitivity induced by pioglitazone treatment is associated with enhanced expression of genes representing ribosomal protein biosynthesis and OXPHOS pathways, the latter possibly mediated by an increased expression of PGC-1α. Furthermore, we provide evidence that treatment with pioglitazone corrects preexisting downregulation of the same pathways in skeletal muscle of patients with PCOS. These findings suggest that the insulin-sensitizing effect of pioglitazone therapy may include reversal of abnormalities in ribosomal protein biosynthesis and mitochondrial oxidative phosphorylation.

In a recent microarray study [16], we found reduced expression of OXPHOS genes in skeletal muscle biopsies obtained from the most insulin resistant women among a larger cohort of PCOS patients. Here, we confirm the downregulation of OXPHOS genes in skeletal muscle of PCOS patients, who were randomly assigned to treatment with pioglitazone in the same study cohort. In addition, we found reduced expression of the ribosomal protein pathway using both GenMAPP and GSEA. Although this abnormality was not detected in our previous study of PCOS patients [16], reduced gene expression of ribosomal proteins has been reported in skeletal muscle [19] and adipose tissue [23] of patients with type 2 diabetes in microarray-based studies. Moreover, GO analysis showed that protein biosynthesis was the most downregulated biological process term. These findings support the hypothesis of an association between reduced protein synthesis and insulin resistance in skeletal muscle, such as shown in age-related sarcopenia [24,25]. However, Halvatsiotis *et al* [26] reported no changes in muscle protein synthesis rate in patients with type 2 diabetes, indicating that transcriptional downregulation of the translational machinery is not necessarily reflected by a decreased protein synthesis rate. Thus, whether muscle protein synthesis rate is affected in PCOS patients remains to be determined, but our data indicate that reduced gene expression of ribosomal proteins and other pathways involved in protein metabolism can not be attributed to hyperglycemia or obesity, and, therefore, may represent an early defect associated with insulin resistance in PCOS.

In the present study, the major finding is that the expression of genes representing OXPHOS pathways is upregulated in skeletal muscle of PCOS patients together with an increase in insulin sensitivity in response to 16 weeks treatment with pioglitazone. To our knowledge, no study has addressed the effects of TZDs on muscle transcripts in PCOS. Similar to other studies using pathway analysis [16,18,19], we observed small, but significant, changes in expression of many genes representing OXPHOS pathways, and some of these changes were also confirmed by q-RT-PCR. Thus, these data provide evidence for an association between small, but coordinated, improvements in expression of OXPHOS genes in skeletal muscle and insulin sensitivity after TZD treatment in women with PCOS. In a recent report, Pagel-Langenickel *et al* [27] showed that improved glycemic control in patients with type 2 diabetes treated with rosiglitazone for 12 weeks was independent of muscle mitochondrial content/activity. Thus, rosiglitazone-mediated improvement in OXPHOS protein content and citrate synthase activity was observed only in individuals with relatively preserved maximal oxygen consumption and mitochondrial copy numbers. Indeed, the difference between our study and the study by Pagel-Langenickel *et al* [27] may reflect a modestly longer duration of therapy (16 vs. 12 weeks), and the broader PPAR isoform activation activity of pioglitazone versus other TZDs such as rosiglitazone. Another possibility is that intervention with TZDs earlier in the spectrum of insulin resistance, e.g. in normoglycemic patients with PCOS, has a greater effect on improving mitochondrial biogenesis than in patients with type 2 diabetes. Moreover, expression of muscle mRNA may not accurately reflect the abundance or activity of OXPHOS proteins, and further studies are warranted to establish

whether improved OXPHOS gene expression in response to TZDs is associated with increased content and function of muscle mitochondria in PCOS.

PPARγ ligands such as TZDs may enhance skeletal muscle insulin sensitivity by inducing expression of genes like adiponectin or genes involved in fatty acid uptake and oxidation in adipose tissue which, in turn, convey signals to skeletal muscle [4,28]. Thus, recombinant adiponectin stimulates fatty acid oxidation and glucose transport by activation of AMPK in rodent muscle [29,30]. Other studies have shown that TZDs also may exert their action independently of PPARγ [31] and adiponectin, e.g. by direct activation of AMPK [32,33]. Chronic activation of AMPK is associated with increased mitochondrial biogenesis [34], probably mediated by increased expression of PGC-1α and NRF-1 [34,35]. In a single study of skeletal muscle from type 2 diabetic patients, Bandyopadhyay *et al* [12] showed that chronic rosiglitazone therapy restored total AMPK activity. Contradictory to this, the activity and protein content of AMPK in muscle were not increased in our pioglitazone treated PCOS subjects, despite a 2-fold increase in circulating adiponectin [36]. This may indicate that other mechanisms independently of AMPK mediated the effect of TZDs on PGC-1α transcription, mitochondrial biogenesis and insulin sensitivity in PCOS, e.g. through calcium/calmodulin-dependent protein kinase and p38 mitogen-activated protein kinase [37,38]. Recently, it was shown that the effects of AMPK on gene expression of glucose transporter 4, mitochondrial genes and PGC-1α, are mediated almost entirely by phosphorylation of PGC-1α at specific sites [39]. Thus, another possibility is that pioglitazone treatment of women with PCOS causes a transient increase in AMPK activity, which then disappears in response to enhanced OXPHOS gene expression and ATP synthesis.

Downregulation of PGC-1α is associated with reduced expression of OXPHOS genes in skeletal muscle of women with PCOS [16], patients with type 2 diabetes [18,19], and their first-degree relatives [19]. In the current study, treatment with pioglitazone increased the expression of PGC-1α in muscle of PCOS patients providing an explanation at the molecular level for the observed upregulation of OXPHOS genes. These data are consistent with a recent study showing improved skeletal muscle oxidative enzyme activity and restoration of PGC-1α gene expression upon rosiglitazone treatment in obese patients with type 2 diabetes [40]. Whether the effect on PGC-1α is mediated by the 2-fold increase in adiponectin, a direct effect of pioglitazone on PPARγ, or other as yet unknown effects of pioglitazone remains to be clarified.

Another mechanism by which pioglitazone may enhance OXPHOS gene expression in PCOS is via the effect of insulin on mitochondrial biogenesis. Thus, studies have shown that acute and chronic insulin infusion enhances the transcriptional activity of genes involved in pathways such as energy metabolism in skeletal muscle of healthy subjects and patients with type 2 diabetes [41–43]. In the present study, a pathway termed Insulin 2F Up, representing genes upregulated 2-fold by insulin stimulation of skeletal muscle in healthy subjects [41], was significantly upregulated in response to pioglitazone. Thus, it is possible that the effect of pioglitazone on OXPHOS gene expression is, at least in part, mediated by an improved insulin action in muscle of PCOS patients. However, further studies are needed to address the exact molecular mechanisms by which pioglitazone treatment improves insulin action on transcriptional activity in women with PCOS.

Using two different approaches for global pathway analysis, we observed a consistent coordinated up-regulation of genes involved in ribosomal protein biosynthesis in muscle of PCOS patients after pioglitazone administration. A similar effect of TZDs on muscle protein metabolism has, to our knowledge, not been reported previously in PCOS or other insulin resistant conditions. Our data indicate that the beneficial effects of pioglitazone therapy in PCOS may include improved protein metabolism mediated either via an enhanced anabolic action of insulin or directly via PPARγ by an as yet unknown mechanism. Several lines of evidence support a role for increased insulin action. Thus, insulin promotes protein synthesis by rapidly activating several components of the translational machinery [44]. This response is elicited primarily through phosphoinositide 3-kinase (PI3K) and Akt, which via inhibition of GSK-3 causes dephosphorylation and activation of eIF2B, and via activation of the mammalian target of rapamycin (mTOR) signaling and phosphorylation of its downstream targets ribosomal S6 kinase 1 (S6K1) and activation of eukaryotic initiation factor 4E binding protein 1 (4EBP1), promotes initiation and elongation [44]. Although our data provide evidence that the increased transcription of ribosomal proteins was not caused by changes in the expression of these key components in insulin signaling to protein synthesis, we have recently reported that impaired insulin action on Akt phosphorylation in muscle of PCOS patients was normalized after treatment with pioglitazone [36]. Thus, changes in the activity of mTOR signaling enzymes and eIF2B provides a possible explanation for the observed changes in the translational machinery before and after pioglitazone treatment. Moreover, recent microarray-based studies of gene expression in skeletal muscle of healthy humans have shown that the majority of the genes upregulated in response to acute insulin infusion codes for proteins involved in transcriptional and translational regulation including a number of ribosomal proteins [41,45]. These reports lend support to the hypothesis that in the long term, insulin also increases the cellular content of ribosomes augmenting the capacity for protein synthesis. Interestingly, there is experimental evidence that mTOR activity may also play a regulatory role for mitochondrial metabolism [46]. Thus, improved expression of genes involved in ribosomal protein biosynthesis and mitochondrial biogenesis in skeletal muscle of PCOS patients after long term pioglitazone therapy may involve enhanced mTOR signaling mediated in part by an improved anabolic action of insulin.

In summary, pioglitazone significantly upregulated a large number of pathways in skeletal muscle of women with PCOS. The most important finding was that pioglitazone increased the expression of genes representing OXPHOS and ribosomal protein biosynthesis, and that this effect corrected preexisting downregulation of genes in the same pathways. Upregulation of OXPHOS genes seems to involve increased expression of PGC-1α, and may occur via an increase in adiponectin, a direct effect of TZD on muscle PPAR gamma, and/or via an improved insulin action on mitochondrial biogenesis. Further studies are required to assess the precise mechanisms by which prolonged treatment with TZD increases expression of OXPHOS and ribosomal protein biosynthesis in skeletal muscle, and how these transcriptional changes improves insulin sensitivity.

Materials and Methods

Subjects

Ten obese Caucasian women of reproductive age with PCOS participated in the study to test the effect of pioglitazone therapy on skeletal muscle gene expression. These subjects represent all women with PCOS from whom a muscle biopsy was obtained both before and after treatment with pioglitazone among a larger cohort of PCOS patients [7,16]. Briefly, thirty PCOS patients were randomly assigned to either pioglitazone or placebo. No

effect with respect to hormonal or metabolic parameters was found in the placebo group after treatment cessation.

Criteria for PCOS included irregular periods with cycle length >35 days during the last year, free testosterone level above reference interval (>0.035 nmol/l), and/or hirsutism (total Ferriman-Gallwey score>7) [7]. All PCOS patients accepted to withdraw oral contraceptives >3 months before evaluation and consented to use barrier contraception combined with spermatocidal cream during the study period. Control subjects had regular menses, normal glucose tolerance, and no family history of diabetes. Women with diabetes (fasting plasma glucose≥7.0 mmol/l), hypertension, elevated liver enzyme levels, adrenal enzyme defects, hyperprolactinemia, and hypothyroidism were excluded from the study. Participants were excluded if they were pregnant. All participants had normal results on screening blood tests of hepatic and renal function. No subjects were taking any medication known to affect hormonal or metabolic parameters. Informed written consent was obtained from all subjects before participation. The study was approved by the Local Ethics Committee and was performed in accordance with the Helsinki Declaration.

Study design

The design of the study has previously been described [7]. All subjects were instructed to refrain from strenuous physical activity for a period of 48-h before the euglycemic-hyperinsulinemic clamp studies. The study subjects were admitted to the Diabetes Research Centre at Odense University Hospital at 08:00 after an overnight fast. To test the effect of a PPAR-γ agonist, pioglitazone (30 mg/day; Actos, Takeda, Lilly A/S, Lyngby, Denmark) or placebo was administered to PCOS patients for a period of 16 weeks. One patient receiving pioglitazone was excluded from the study because of side effects (dizziness, ankle edema, and anxiety). After the treatment period, the initial evaluation program was repeated [7].

Rates of total glucose disposal rates (Rd), glucose and lipid oxidation, and non-oxidative glucose disposal (NOGD) were assessed by euglycemic-hyperinsulinemic clamp studies (4-h insulin infusion, 40 mU7min per m^2) combined with indirect calorimetry as described in detail previously [7]. Skeletal muscle biopsies were obtained in the basal steady-state period of the clamp from the vastus lateralis muscle using a modified Bergström needle with suction under local anaesthesia (10–15 ml of lidocain 2% (20 g/l)). Muscle biopsies were immediately blotted free of blood, fat, and connective tissue and frozen in liquid nitrogen within 30 s. Assays for serum levels of insulin, free testosterone, luteinizing hormone (LH), follicle-stimulating hormone (FSH), plasma glucose, triglyceride, and free fatty acids (FFA) were as described in [7].

RNA extraction and microarray preparation

Total RNA was extracted from skeletal muscle tissue using the TRIzol protocol (Life Technologies, Gaithersburg, MD). An extra phenol-chloroform step was included as described previously [16]. Quantity of RNA was determined with a spectrophotometer, and RNA of high quality was assessed using Agilent 2100 Bioanalyser (Agilent Technologies, Palo Alto, CA) and degradometer software [47]. The MessageAmpTM II-Biotin single round aRNA amplification kit (Ambion, Austin, TX) was applied to convert one μg of purified total RNA to biotin-labeled aRNA. Labeled aRNA was fragmented as described in the Affymetrix manual (Affymetrix, Santa Clara, CA) and hybridized to Affymetrix HG-U133 Plus 2.0 chips. The 3′/5′ ratio of Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was below 1.42 for all chips and confirmed high quality RNA.

Data processing

The R statistical software [48] was applied for data preprocessing (www.bioconductor.org) and statistical analysis. Global background correction of probe intensities was done by employing a method implemented in the robust multi-array average (RMA) method [49], and data was normalized by using constant normalization. We tried to use variance stabilizing transformation (VSN) for normalization as well. The results were comparable to those obtained using constant normalization (data not shown). Gene expression index calculation was done using model based index calculation (MBEI) by Li & Wong [50]. Only perfect match probes were included in data analysis. Differences in gene expression between PCOS patients before and after pioglitazone therapy and between PCOS patients and controls were calculated for each gene by using Welch two sample t-test for paired and unpaired data, respectively. An uncorrected $P<0.05$ was considered significant. Data are available from GEO (http://www.ncbi.nlm.nih.gov/geo, Accession No. GSE8157).

Global pathway analysis

GenMAPP 2.1 [51] (MAPPFinder 2.1 (20)), and GSEA 2.0.1 [52] were employed to assess significantly regulated pathways, gene ontology (GO) terms, and gene sets in the two data sets. In MAPPFinder, a total of 203 pathways and 6251 GO terms were applied whereas 189 gene sets were used in GSEA. An enrichment score (ES) was calculated for each gene set in GSEA and the statistical significance of the ES was estimated by an empirical permutation test using 2.000 gene permutations to obtain the nominal p-value.

Quantitative realtime PCR (Q-RT-PCR)

Total RNA from 9 PCOS patients before and after treatment with pioglitazone was treated with DNAse I (USB, Cleveland, Ohio) and reverse transcribed to single-stranded cDNA using TaqMan reverse transcription reagents and random hexamer primers (Applied Biosystems). One patient was not included in the analysis because of a lack of RNA. TaqMan gene expression assays (Applied Biosystems) and TaqMan Universal Master Mix (Applied Biosystems) were used to quantify gene expression changes using Applied Biosystems Prism 7700 (Supplemental Table S9). β-actin (Applied Biosystems) was used as a housekeeping gene to normalize gene expression levels. To perform the most appropriate validation of microarray data, bioinformatic approaches such as NetAffx (www.Affymetrix.com), refseq (www.ncbi.nlm.nih.gov), and Ensembl (www.ensembl.org) were used to identify the TaqMan probe sequence for each gene, which has the highest similarity to the Affymetrix probe set (Supplemental Table S9).

Supporting Information

Table S1 The ten most downregulated pathways analyzed with MAPPFinder 2.1.
Found at: doi:10.1371/journal.pone.0002466.s001 (0.05 MB DOC)

Table S2 Ranking of the ten most downregulated gene sets analyzed with GSEA 2.0.1.
Found at: doi:10.1371/journal.pone.0002466.s002 (0.05 MB DOC)

Table S3 The twenty most upregulated GO terms analyzed with MAPPFinder 2.1.
Found at: doi:10.1371/journal.pone.0002466.s003 (0.07 MB DOC)

Table S4 Ranking of the twenty most downregulated GO terms analyzed with MAPPFinder 2.1.
Found at: doi:10.1371/journal.pone.0002466.s004 (0.07 MB DOC)

Table S5 Ranking of the ten most upregulated pathways analyzed with MAPPFinder 2.1.
Found at: doi:10.1371/journal.pone.0002466.s005 (0.05 MB DOC)

Table S6 The top-ten upregulated gene sets analyzed with GSEA 2.0.1.
Found at: doi:10.1371/journal.pone.0002466.s006 (0.05 MB DOC)

Table S7 The twenty most upregulated GO terms analyzed with MAPPFinder 2.1.
Found at: doi:10.1371/journal.pone.0002466.s007 (0.07 MB DOC)

Table S8 Ranking of the twenty most downregulated GO terms analyzed with MAPPFinder 2.1.
Found at: doi:10.1371/journal.pone.0002466.s008 (0.07 MB DOC)

Table S9 Probe set and TaqMan assay for the 18 selected genes.
Found at: doi:10.1371/journal.pone.0002466.s009 (0.08 MB DOC)

Author Contributions

Conceived and designed the experiments: KH VS SK TK HB DG. Performed the experiments: VS DG. Analyzed the data: KH VS SK QT TJ DG. Contributed reagents/materials/analysis tools: VS TK HB. Wrote the paper: KH VS SK QT TJ TK HB DG. Other: Contributed to the final version: TK TJ DG QT SK HB. Wrote the initial draft of the paper: VS KH.

References

1. Ehrmann DA (2005) Polycystic ovary syndrome. N Engl J Med 352: 1223–1236.
2. Dunaif A (1997) Insulin resistance and the polycystic ovary syndrome: mechanism and implications for pathogenesis. Endocr Rev 18: 774–800.
3. Yki-Jarvinen H (2004) Thiazolidinediones. N Engl J Med 351: 1106–1118.
4. Olefsky JM, Saltiel AR (2000) PPAR gamma and the treatment of insulin resistance. Trends Endocrinol Metab 11: 362–368.
5. Feinstein DL, Spagnolo A, Akar C, Weinberg G, Murphy P, et al. (2005) Receptor-independent actions of PPAR thiazolidinedione agonists: is mitochondrial function the key? Biochem Pharmacol 70: 177–188.
6. Loviscach M, Rehman N, Carter L, Mudaliar S, Mohadeen P, et al. (2000) Distribution of peroxisome proliferator-activated receptors (PPARs) in human skeletal muscle and adipose tissue: relation to insulin action. Diabetologia 43: 304–311.
7. Glintborg D, Hermann AP, Andersen M, Hagen C, Beck-Nielsen H, et al. (2006) Effect of pioglitazone on glucose metabolism and luteinizing hormone secretion in women with polycystic ovary syndrome. Fertil Steril 86: 385–397.
8. Brettenthaler N, De GC, Huber PR, Keller U (2004) Effect of the insulin sensitizer pioglitazone on insulin resistance, hyperandrogenism, and ovulatory dysfunction in women with polycystic ovary syndrome. J Clin Endocrinol Metab 89: 3835–3840.
9. Miyazaki Y, Mahankali A, Matsuda M, Glass L, Mahankali S, et al. (2001) Improved glycemic control and enhanced insulin sensitivity in type 2 diabetic subjects treated with pioglitazone. Diabetes Care 24: 710–719.
10. Miyazaki Y, He H, Mandarino LJ, DeFronzo RA (2003) Rosiglitazone improves downstream insulin receptor signaling in type 2 diabetic patients. Diabetes 52: 1943–1950.
11. Kim YB, Ciaraldi TP, Kong A, Kim D, Chu N, et al. (2002) Troglitazone but not metformin restores insulin-stimulated phosphoinositide 3-kinase activity and increases p110beta protein levels in skeletal muscle of type 2 diabetic subjects. Diabetes 51: 443–448.
12. Bandyopadhyay GK, Yu JG, Ofrecio J, Olefsky JM (2006) Increased malonyl-CoA levels in muscle from obese and type 2 diabetic subjects lead to decreased fatty acid oxidation and increased lipogenesis; thiazolidinedione treatment reverses these defects. Diabetes 55: 2277–2285.
13. Wilmsen HM, Ciaraldi TP, Carter L, Reehman N, Mudaliar SR, Henry RR (2003) Thiazolidinediones upregulate impaired fatty acid uptake in skeletal muscle of type 2 diabetic subjects. Am J Physiol Endocrinol Metab 285: E354–E362.
14. Lessard SJ, Chen ZP, Watt MJ, Hashem M, Reid JJ, et al. (2006) Chronic rosiglitazone treatment restores AMPKalpha2 activity in insulin-resistant rat skeletal muscle. Am J Physiol Endocrinol Metab 290: E251–E257.
15. Jove M, Salla J, Planavila A, Cabrero A, Michalik L, et al. (2004) Impaired expression of NADH dehydrogenase subunit 1 and PPARgamma coactivator-1 in skeletal muscle of ZDF rats: restoration by troglitazone. J Lipid Res 45: 113–123.
16. Skov V, Glintborg D, Knudsen S, Jensen T, Kruse TA, et al. (2007) Reduced expression of nuclear-encoded genes involved in mitochondrial oxidative metabolism in skeletal muscle of insulin-resistant women with polycystic ovary syndrome. Diabetes 56: 2349–2355.
17. Lockhart DJ, Dong H, Byrne MC, Follettie MT, Gallo MV, et al. (1996) Expression monitoring by hybridization to high-density oligonucleotide arrays. Nat Biotechnol 14: 1675–1680.
18. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, et al. (2003) PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet 34: 267–273.
19. Patti ME, Butte AJ, Crunkhorn S, Cusi K, Berria R, et al. (2003) Coordinated reduction of genes of oxidative metabolism in humans with insulin resistance and diabetes: Potential role of PGC1 and NRF1. Proc Natl Acad Sci U S A 100: 8466–8471.
20. Doniger SW, Salomonis N, Dahlquist KD, Vranizan K, Lawlor SC, Conklin BR (2003) MAPPFinder: using Gene Ontology and GenMAPP to create a global gene-expression profile from microarray data. Genome Biol 4: R7.
21. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Statist Soc B 57: 289–300.
22. GSEA: [www broad mit edu/gsea/msigdb/cards /c2_cards_ index html] 2007.
23. Dahlman I, Forsgren M, Sjogren A, Nordstrom EA, Kaaman M, et al. (2006) Downregulation of Electron Transport Chain Genes in Visceral Adipose Tissue in Type 2 Diabetes Independent of Obesity and Possibly Involving Tumor Necrosis Factor-{alpha}. Diabetes 55: 1792–1799.
24. Barazzoni R (2004) Skeletal muscle mitochondrial protein metabolism and function in ageing and type 2 diabetes. Curr Opin Clin Nutr Metab Care 7: 97–102.
25. Rasmussen BB, Fujita S, Wolfe RR, Mittendorfer B, Roy M, et al. (2006) Insulin resistance of muscle protein metabolism in aging. FASEB J 20: 768–769.
26. Halvatsiotis P, Short KR, Bigelow M, Nair KS (2002) Synthesis rate of muscle proteins, muscle functions, and amino acid kinetics in type 2 diabetes. Diabetes 51: 2395–2404.
27. Pagel-Langenickel I, Schwartz DR, Arena RA, Minerbi DC, Johnson DT, et al. (2007) A discordance in rosiglitazone mediated insulin sensitization and skeletal muscle mitochondrial content/activity in Type 2 diabetes mellitus. Am J Physiol Heart Circ Physiol 293: H2659–2666.
28. Pasquali R, Gambineri A (2006) Insulin-sensitizing agents in polycystic ovary syndrome. Eur J Endocrinol 154: 763–775.
29. Yamauchi T, Kamon J, Minokoshi Y, Ito Y, Waki H, et al. (2002) Adiponectin stimulates glucose utilization and fatty-acid oxidation by activating AMP-activated protein kinase. Nat Med 8: 1288–1295.
30. Tomas E, Tsao TS, Saha AK, Murrey HE, Zhang CC, et al. (2002) Enhanced muscle fat oxidation and glucose transport by ACRP30 globular domain: acetyl-CoA carboxylase inhibition and AMP-activated protein kinase activation. Proc Natl Acad Sci U S A 99: 16309–16313.
31. Brunmair B, Gras F, Neschen S, Roden M, Wagner L, et al. (2001) Direct thiazolidinedione action on isolated rat skeletal muscle: fuel handling is independent of peroxisome proliferator-activated receptor-gamma-mediated changes in gene expression. Diabetes 50: 2309–2315.
32. Lebrasseur NK, Kelly M, Tsao TS, Farmer SR, Saha AK, et al. (2006) Thiazolidinediones can rapidly activate AMP-activated protein kinase in mammalian tissues. Am J Physiol Endocrinol Metab 291: E175–E181.
33. Hardie DG (2007) AMP-activated protein kinase as a drug target. Annu Rev Pharmacol Toxicol 47: 185–210.
34. Bergeron R, Ren JM, Cadman KS, Moore IK, Perret P, et al. (2001) Chronic activation of AMP kinase results in NRF-1 activation and mitochondrial biogenesis. Am J Physiol Endocrinol Metab 281: E1340–E1346.
35. Zong H, Ren JM, Young LH, Pypaert M, Mu J, et al. (2002) AMP kinase is required for mitochondrial biogenesis in skeletal muscle in response to chronic energy deprivation. Proc Natl Acad Sci U S A 99: 15983–15987.
36. Højlund K, Glintborg D, Andersen NR, Birk JB, Treebak JT, et al. (2008) Impaired insulin-stimulated phosphorylation of Akt and AS160 in skeletal muscle of women with polycystic ovary syndrome is reversed by pioglitazone treatment. Diabetes 57: 357–366.
37. Wu H, Kanatous SB, Thurmond FA, Gallardo T, Isotani E, et al. (2002) Regulation of mitochondrial biogenesis in skeletal muscle by CaMK. Science 296: 349–352.
38. Wright DC, Geiger PC, Han DH, Jones TE, Holloszy JO (2007) Calcium induces increases in peroxisome proliferator-activated receptor gamma coacti-

vator-1alpha and mitochondrial biogenesis by a pathway leading to p38 mitogen-activated protein kinase activation. J Biol Chem 282: 18793–18799.
39. Jager S, Handschin C, St-Pierre J, Spiegelman BM (2007) AMP-activated protein kinase (AMPK) action in skeletal muscle via direct phosphorylation of PGC-1alpha. Proc Natl Acad Sci U S A 104: 12017–12022.
40. Mensink M, Hesselink MK, Russell AP, Schaart G, Sels JP, Schrauwen P (2007) Improved skeletal muscle oxidative enzyme activity and restoration of PGC-1alpha and PPARbeta/delta gene expression upon rosiglitazone treatment in obese patients with type 2 diabetes mellitus. Int J Obes (Lond) 31: 1302–1310.
41. Rome S, Clement K, Rabasa-Lhoret R, Loizon E, Poitou C, et al. (2003) Microarray profiling of human skeletal muscle reveals that insulin regulates approximately 800 genes during a hyperinsulinemic clamp. J Biol Chem 278: 18063–18068.
42. Sreekumar R, Halvatsiotis P, Schimke JC, Nair KS (2002) Gene expression profile in skeletal muscle of type 2 diabetes and the effect of insulin treatment. Diabetes 51: 1913–1920.
43. Stump CS, Short KR, Bigelow ML, Schimke JM, Nair KS (2003) Effect of insulin on human skeletal muscle mitochondrial ATP production, protein synthesis, and mRNA transcripts. Proc Natl Acad Sci U S A 100: 7996–8001.
44. Proud CG (2006) Regulation of protein synthesis by insulin. Biochem Soc Trans 34: 213–216.
45. Wu X, Wang J, Cui X, Maianu L, Rhees B, et al. (2007) The effect of insulin on expression of genes and biochemical pathways in human skeletal muscle. Endocrine 31: 5–17.
46. Schieke SM, Phillips D, McCoy JP Jr, Aponte AM, Shen RF, et al. (2006) The mammalian target of rapamycin (mTOR) pathway regulates mitochondrial oxygen consumption and oxidative capacity. J Biol Chem 281: 27643–27652.
47. Auer H, Lyianarachchi S, Newsom D, Klisovic MI, Marcucci G, Kornacker K (2003) Chipping away at the chip bias: RNA degradation in microarray analysis. Nat Genet 35: 292–293.
48. Ihaka R, Gentleman R (1996) A language for data analysis and graphics. J Comp Graph Stat 5: 299–314.
49. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, et al. (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 4: 249–264.
50. Li C, Wong WH (2001) Model-based analysis of oligonucleotide arrays: model validation, design issues and standard error application. Genome Biol 2: 1–11.
51. Dahlquist KD, Salomonis N, Vranizan K, Lawlor SC, Conklin BR (2002) GenMAPP, a new tool for viewing and analyzing microarray data on biological pathways. Nat Genet 31: 19–20.
52. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, et al. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 102: 15545–15550.